Principia Logico-Metaphysica
(Draft/Excerpt)

Edward N. Zalta
Philosophy Department
Stanford University

With critical theoretical contributions by

Daniel Kirchner
Institut für Mathematik
Freie Universität Berlin
and
Uri Nodelman
Philosophy Department
Stanford University

June 4, 2023

https://mally.stanford.edu/principia.pdf

This document was typeset with kpfonts using the LaTeX document preparation system.

Library of Congress Cataloging-in-Publication Data

Zalta, Edward N., 1952–
Principia Logico-Metaphysica

Bibliography: p.
Doesn't yet include index.
1. Logic, Symbolic 2. Metaphysics (Philosophy). I. Title

Catalog No. xxxxxx 2023 xxx-xxxx xx-xxxxx
ISBN X-XXX-XXXXX-X

To my wife, Susanne Z. Riehemann

Draft / Excerpt

NOTE: This is an excerpt from an incomplete draft of the monograph *Principia Logica-Metaphysica*. The monograph draft currently has four parts:

Part I: Prophilosophy
Part II: Philosophy
Part III: Metaphilosophy
Part IV: Technical Appendices, Bibliography, Index

This excerpt was generated on June 4, 2023 and contains:

- Part II:

- Part IV:

Consequently, this excerpt omits the Preface, Part I, Part III (which is mostly unwritten), and some Appendices in Part IV. The present excerpt sometimes contains references to the omitted content and *active links in the Table of Contents to omitted content won't work*.

The work is ongoing and so the monograph changes constantly. Any citations to this material should explicitly reference this version of June 4, 2023, since page numbers, chapter numbers, section numbers, item (definition, theorem) numbers, etc., may all change in future versions.

Contents

Acknowledgments

Object theory would not be where it is today without the help of numerous colleagues. I've coauthored papers with many of these colleagues, had enlightening discussions with others, and received email messages with suggested corrections from still others.

Daniel Kirchner and Uri Nodelman deserve special mention; they are listed on the title page as having made critical theoretical contributions. This is no understatement. Though I authored the *monograph* (with the exception of Chapter 14, which was coauthored with Nodelman), both Kirchner and Nodelman coauthored the *axiomatic theory* described in the monograph. I'd also like to acknowledge Daniel West, who read early drafts of nearly all of Part II, and contributed a number of interesting observations, proofs, and even a helpful metarule, all of which constituted improvements. I'm also grateful to Wes Anderson, who also read significant parts of Part II and notified me of a number of errors.

Until this monograph gets into a state closer to publication, I am simply going to list the names of my other colleagues who've contributed in some way to this monograph, in alphabetical order: Peter Aczel, Jesse Alama, Colin Allen, C. Anthony Anderson, David James Anderson, Guillermo Badía Hernández, Christoph Benzmüller, Johannes Brandl, Otávio Bueno, Mark Colyvan, Kit Fine, Branden Fitelson, Richard Grandy, Allen Hazen, Alexander Hieke, Michael Jubien, Jeffrey Kegler, Johannes Korbmacher, Fred Kroon, Hannes Leitgeb, David Lewis, Godehard Link, Bernard Linsky, Alan McMichael, Christopher Menzel, Edgar Morscher, Michael Nelson, Karl-Georg Niebergall, Paul E. Oppenheimer, Terence Parsons, Barbara Partee, Francis Jeffrey Pelletier, Gideon Rosen, Tony Roy, Dana Scott, Merel Semeijn, Peter Simons, Chris Swoyer, Johan van Benthem, and Kai Wehmeier.

Finally, and most importantly, I'd like to thank my wife, Susanne Z. Riehemann, who has joined me in so many adventures, both intellectual and in the outdoors, and brought me such joy and happiness. I extremely grateful to her for sharing her life and making it possible for me to write this book. I've been lucky to have a life partner who is keenly interested in my work, and whose own fields of study and investigation so nicely dovetail with my own.

Stanford, CA and Mountain View, CA
June 4, 2023

Part I

Prophilosophy

Part II

Philosophy

Chapter 7

The Language

Throughout most of Part II, we prove theorems that are formulated in a second-order modal language. This will be the language used in Chapters 7 – 14. However this language is just a fragment of a more general, type-theoretic modal language. We postpone the definition of the type-theoretic language until Chapter 15, where we investigate typed object theory and its applications.

In the present and subsequent chapters, our *metalanguage* makes use of some basic notions and principles of number theory and set theory, so as to more precisely articulate certain definitions. But *none* of these notions and principles are used in the object language defined in this chapter. Ultimately, the philosophical system sketched over the next few chapters will offer us an analysis of the basic notions and principles of number theory and set theory used in the metalanguage, but we won't be in a position to see this until Chapter 10 (where we define and prove facts about *natural classes*) and Chapter 14 (where we define and prove facts about *natural cardinals* and *natural numbers*).

7.1 Metatheoretical Definitions

The definitions that generate a particular second-order language are given over the next several items.

(**1**) **Metadefinitions:** Simple Terms. A *simple term* of our second-order language is any expression that is a *simple individual term* or a *simple n-place relation term* ($n \geq 0$), where these are listed as follows:

(.1) Simple Individual Terms: (Less Formal)

Individual Constants (Names):

$a_1, a_2, \ldots$ $(a, b, c, b_1, b_2, \ldots, c_1, c_2, \ldots)$

Individual Variables:
$x_1, x_2, \ldots$ $(x, y, z, u, v, w, y_1, y_2, \ldots, z_1, z_2, \ldots)$

(.2) Simple n-ary Relation Terms ($n \geq 0$):
n-ary Relation Constants (Names)
$P^n_1, P^n_2, \ldots$ $(P^n, Q^n, R^n, S^n, T^n)$
n-ary Relation Variables:
$F^n_1, F^n_2, \ldots$ $(F^n, G^n, H^n, I^n, J^n)$

(.3) Distinguished Unary Relation Constant:
$E!$ (read: 'being concrete' or 'concreteness')
where $E!$ is just a rewritten version of the unary relation constant P^1_1

In what follows, we shall use the technical term *primitive constant* as follows:

(.4) A *primitive constant* is any simple individual constant or simple n-ary relation constant (for some $n \geq 0$) that occurs in the lists in (.1), (.2), and (.3) above.

This helps us to distinguish primitive constants from *new* constants introduced by definition. No similar distinction is needed for variables, since we won't introduce new variables into the language by definition. To facilitate readability, we often use the expressions listed in the column labeled 'Less Formal' as replacements for the official expressions of the language.

(2) **Metadefinitions:** Syncategorematic Expressions. A syncategorematic expression represents a primitive notion of the language but is neither a term (i.e., the kind of expression that may have a denotation) nor a formula (i.e., the kind of expression that has truth conditions). To list the syncategorematic expressions of our language, we use α as a metavariable that ranges over all variables and use ν (Greek *nu*), sometimes decorated with a numerical subscript, as a metavariable that ranges just over individual variables:

(.1) Unary Formula-Forming Operators:
$\neg$ ('it is not the case that' or 'it is false that')
$\square$ ('necessarily' or 'it is necessary that')
$\mathcal{A}$ ('actually' or 'it is actually the case that')

(.2) Binary Formula-Forming Operator:
$\rightarrow$ ('if ..., then ...')

(.3) Variable-Binding Formula-Forming Operator:
$\forall\alpha$ ('every α is such that')
for every variable α

(.4) Variable-Binding Individual-Term-Forming Operator:
$\iota\nu$ ('the ν such that')
for every individual variable ν

(.5) Variable-Binding n-ary Relation-Term-Forming Operators ($n \geq 0$):
$\lambda\nu_1 \ldots \nu_n$ ('being $\nu_1, \ldots, \nu_n$ such that')
for any distinct individual variables $\nu_1, \ldots, \nu_n$, and
λ ('that ...')
where no variables follow the λ.

These primitive, syncategorematic expressions are referenced in the definition of the syntax of our language and are used to define complex formulas and complex terms. In what follows, we sometimes call:

$\neg$	the *negation* operator
$\Box$	the *necessity* operator
$\mathcal{A}$	the *actuality* operator
$\rightarrow$	the *conditional* operator
$\forall$	the *universal quantifier*
ι	the *definite description* operator
λ	the *relation abstraction*, or λ, operator

By convention, in any conditional formula of the form $\varphi \rightarrow \psi$, we say φ is the *antecedent* and ψ the *consequent* of the conditional.

(3) **Metadefinitions:** Syntax of the 'Second-order' Language. We present the syntax of our second-order language by a simultaneous recursive definition of the following four kinds of expressions: *individual term*, *n-place relation term*, *formula*, and *term*.

Base Clauses:

(.1) Every simple individual term (i.e., every individual constant and individual variable) is an individual term and every simple n-ary relation term (i.e., every n-ary relation constant and n-ary relation variable), is an n-ary relation term ($n \geq 0$)

(.2) Every 0-ary relation constant and 0-ary relation variable is a formula

(.3) If Π^n is any n-ary relation term ($n \geq 1$) and $\kappa_1, \ldots, \kappa_n$ are any individual terms, then

(.a) $\Pi^n\kappa_1 \ldots \kappa_n$ is a formula ('$\kappa_1, \ldots, \kappa_n$ exemplify Π^n')

(.b) $\kappa_1 \ldots \kappa_n\Pi^n$ is a formula ('$\kappa_1, \ldots, \kappa_n$ encode Π^n')

Recursive Clauses:

(.4) If φ and ψ are formulas and α any variable, then
$[\lambda\, \varphi]$, $(\neg\varphi)$, $(\varphi \rightarrow \psi)$, $\forall\alpha\varphi$, $(\Box\varphi)$, and $(\mathcal{A}\varphi)$ are formulas.

(.5) If φ is any formula and ν any individual variable, then
$\imath\nu\varphi$ is an individual term

(.6) If φ is any formula and $\nu_1,\ldots,\nu_n$ are any distinct individual variables ($n \geq 0$), then

(a) $[\lambda\nu_1 \ldots \nu_n\, \varphi]$ is an n-ary relation term, and

(b) φ itself is a 0-ary relation term.

Finally, we say:

(.7) A *term* is any individual term or n-ary relation term ($n \geq 0$).

Though it should be clear how to read the formulas and terms of the language we've just defined, there are two interesting facts to note about reading certain expressions, namely, (i) 0-ary λ-expressions of the form $[\lambda\varphi]$ are both formulas and terms, and (ii) every formula φ is a 0-ary relation term. We discuss these in turn.

By (.4), expressions of the form $[\lambda\, \varphi]$ are formulas and, by (.6.a), they are also 0-ary relation terms. So expressions of the form $[\lambda\varphi]$ should be read in one of two ways, depending on the context. On the one hand, when $[\lambda\varphi]$ stands by itself or occurs in formula position (e.g., on one side of a conditional or biconditional), it asserts *that φ is true*, since truth is the 0-ary case of predication. So, we read the formula $[\lambda\, \varphi] \equiv \varphi$ as: that-φ is true if and only if φ. On the other hand, there will be contexts in which $[\lambda\, \varphi]$ functions as a term, and in those contexts, we read $[\lambda\, \varphi]$ simply as '*that*-φ'. So, for example, when we define the notion τ *exists* (where τ is any term) and represent it as $\tau\downarrow$, then $[\lambda\, \varphi]\downarrow$ asserts the existence of the proposition $[\lambda\, \varphi]$ and would be read as: *that*-φ *exists*. And when we define the notion τ is identical to σ and represent it a $\tau = \sigma$, then the formula $p = [\lambda\varphi]$ would be read: p is identical to *that*-φ. And when we define the notion *Contingent*(p), where p can be instantiated by any 0-ary relation term, the formula *Contingent*$([\lambda\, \varphi])$ would be read: *that*-φ is contingent. If it helps, one can always preface 'the proposition' to the reading of $[\lambda\varphi]$ when the latter is being used as a term. For instance, one could read our last example, *Contingent*$([\lambda\, \varphi])$, as: the proposition *that*-φ is contingent.

The second interesting fact about reading the language is that (.6.b) stipulates that every formula is a 0-ary relation term. This feature allows us to regard formulas of the form $\varphi \equiv \varphi$ as *substitution instances* of the universal claim $\forall p(p \equiv p)$. But it also means that formulas can occur in contexts where they are simply naming rather than asserting propositions. For example, in the previous paragraph, we noted that $\tau\downarrow$ will be defined for every term. So when a 0-ary relation term such as Fx is substituted for τ, the resulting formula, $Fx\downarrow$, is well-formed. This claim asserts existence of the proposition Fx, in which x is some fixed but unspecified object and F is some fixed but unspecified property. So in the claim $Fx\downarrow$, the formula Fx is not asserting that x exemplifies F, but rather naming the proposition *x-exemplifies-F*. Again, if it helps, one can add 'the proposition' to the reading of a formula being used as term, so that $Fx\downarrow$ may be read: the proposition Fx exists. Similarly, identity will be defined for all terms, including 0-ary relation terms. So a claim of the form $\varphi = \psi$ is well-formed. In claims of this form, the instances of φ and ψ are not making assertions. If it helps, read such claims as: the proposition φ is identical to the proposition ψ. For example, $Fx = [\lambda z\, Fz]x$ would be read: the proposition x exemplifies F is identical to the proposition x exemplifies being an object z that exemplifies F. It should always be clear, in what follows, when a formula is being used, and is to be read, as a term denoting a proposition instead of as an expression making an assertion.

When $n \geq 1$, we call formulas of the form $\Pi^n \kappa_1 \ldots \kappa_n$ (atomic) *exemplification* formulas and formulas of the form $\kappa_1 \ldots \kappa_n \Pi^n$ (atomic) *encoding* formulas. In what follows, we say:

(.8) φ is the *matrix* of $\forall\alpha\varphi$, $\imath\nu\varphi$, and $[\lambda\nu_1 \ldots \nu_n\, \varphi]$ $(n \geq 0)$.

Once we define $\exists\alpha\varphi$, then one should extend the above definition so that φ counts as its matrix.

According to clause (.6.a), any formula φ can serve as the matrix of a relation term of the form $[\lambda\nu_1 \ldots \nu_n\, \varphi]$, where $n \geq 0$. If we allow ourselves to speak informally about the denotation of a term, then it is important to alert the reader to the following facts about the system we shall be developing:

- Not every λ-expression is guaranteed to have a denotation (indeed, some λ-expressions will provably fail to have denotations), and so these expressions, like definite descriptions, will be governed by a negative free logic.

- Every 0-ary λ-expression $[\lambda\, \varphi]$ is guaranteed to have a denotation, by axiom (39.2), and every formula φ is guaranteed to have a denotation, by theorem (104.2).

- It will be provable that $[\lambda\varphi]$ and φ always denote the same 0-ary relation, by theorem (111.1).

In general, λ-expressions are *not* to be interpreted as terms that potentially denote functions, but rather as terms that potentially denote relations. Thus, when we introduce axioms and rules of inference governing λ-expressions in the next two chapters, the resulting λ-calculus is to be understood as a calculus of relations.

Finally, in what follows, we use τ to range over terms. The simple terms listed in (1) are terms in virtue of clause (.1). We say:

(.9) A term τ is *complex* if and only if τ is not a simple term.

So the constants and variables listed in (1) are not complex terms. Given (.9), clauses (.5), and (.6) introduce kinds of complex terms:

- definite descriptions are *complex individual terms*, by (.5) and (.9)
- λ-expressions are *complex relation terms*, by (.6.a) and (.9)
- formulas other than 0-ary relation constants and 0-ary variables are *complex 0-ary relation terms*, by (.6.b) and (.9)

We sometimes also use σ and ρ in addition to τ to range over terms.

(4) **Metadefinition:** A BNF Definition of the Syntax. We may succintly summarize the essential definitions of the context-free grammar of our language using Backus-Naur Form (BNF). In the BNF definition, we repurpose our Greek metavariables as the names of grammatical categories, as follows:

δ	primitive individual constants
ν	individual variables
Σ^n	primitive n-ary relation constants ($n \geq 0$)
Ω^n	n-ary relation variables ($n \geq 0$)
α	variables
κ	individual terms
Π^n	n-ary relation terms ($n \geq 0$)
φ	formulas
τ	terms

The BNF grammar for our second-order language can now be stated as follows:[75]

[75] I'm indebted to Uri Nodelman and Daniel Kirchner, each of whom suggested a different way of simplifying the original BNF definition. Nodelman's suggestion yielded a BNF that could be more efficiently parsed. Kirchner's suggestion led to the elimination of the syntactic category *propositional formula*, i.e., formulas that have encoding subformulas. We may now build well-formed complex n-ary relation terms ($n \geq 1$) of the form $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ from any formula φ, though not all such terms are guaranteed to have a denotation.

$$
\begin{array}{rrl}
 & \delta & ::= a_1, a_2, \ldots \\
 & \nu & ::= x_1, x_2, \ldots \\
(n \geq 0) & \Sigma^n & ::= P^n_1, P^n_2, \ldots \quad \text{(with } P^1_1 \text{ distinguished and written as } E!\text{)} \\
(n \geq 0) & \Omega^n & ::= F^n_1, F^n_2, \ldots \\
 & \alpha & ::= \nu \mid \Omega^n \ (n \geq 0) \\
 & \kappa & ::= \delta \mid \nu \mid \imath\nu\varphi \\
(n \geq 1) & \Pi^n & ::= \Sigma^n \mid \Omega^n \mid [\lambda\nu_1 \ldots \nu_n \ \varphi] \quad (\nu_1, \ldots, \nu_n \text{ are pairwise distinct}) \\
 & \varphi & ::= \Sigma^0 \mid \Omega^0 \mid \Pi^n\kappa_1 \ldots \kappa_n \ (n \geq 1) \mid \kappa_1 \ldots \kappa_n\Pi^n \ (n \geq 1) \mid \\
 & & \qquad [\lambda\, \varphi] \mid (\neg\varphi) \mid (\varphi \rightarrow \varphi) \mid \forall\alpha\varphi \mid (\Box\varphi) \mid (\mathcal{A}\varphi) \\
 & \Pi^0 & ::= \varphi \\
 & \tau & ::= \kappa \mid \Pi^n \quad (n \geq 0)
\end{array}
$$

It is insightful to recognize that one could, in the above BNF, replace φ everywhere by Π^0, except on the next-to-last line, which would then be reversed to read $\varphi ::= \Pi^0$. This alternative BNF would first introduce the individual and relation terms of every arity and then define the formulas as 0-ary relation terms. However, we've made a considered choice to write the BNF as in the above display, not only for ease of readability and understanding, but also because historically, when defining a formal language, the notion of a formula (i.e., an expression that is assertible, has truth conditions, and has logical consequences) are at least as important as terms. The logic developed below should be conceived as governing inferences among formulas rather than a term logic.

Since the penultimate line of the BNF implies that the 0-ary relation terms are precisely the formulas, the metavariables φ and Π^0 range over the same expressions. Usually, it will be more natural to use one metavariable rather than the other to describe some feature of our system. For example, the notion of *subformula* will be defined with respect to formulas φ rather than with respect to 0-ary relation terms Π^0. The definition, see (6) below, implies that φ and ψ are subformulas of the formula $\varphi \rightarrow \psi$. But the formula $p = q$, which is defined in (23.4), will be an instance of the form $\Pi = \Pi'$ (where Π and Π' are 0-ary relation metavariables), since both p and q are being used as 0-ary relation terms. In general, we'll use φ to represent expressions being used to make assertions (and so have truth conditions) and Π^0 to represent expressions being used to denote propositions.

If one defines a finite instance of our language by giving a limiting value to n and listing a finite vocabulary of simple terms, the formulas and terms of the resulting grammar can be parsed by any appropriately-configured off-the-shelf parsing engine using the above BNF.

(5) **Remark:** Notational Conventions. We adopt the following conventions to facilitate readability:

(.1) We often use the less formal expressions listed in (1) instead of their more formal counterparts, and we often drop the superscript indicating the arity of a simple relation term when such terms appear in a formula, since their arity can always be inferred from the number of individual terms in the formula. Thus, instead of writing $P_1^1 a_1$ and $a_2 P^2$, we might write Pa and bQ, respectively; instead of $F_1^1 x_1$ and $x_2 F_2^1$, we might write Fx and yG, respectively; instead of $F_1^2 a_2 x_3$ and $a_3 x_2 F_2^2$, we might write Rbz and cyS; etc. Indeed, we may sometimes use a mixture of formal and less formal expressions in the same formula.

(.2) We substitute $p, q, r, \ldots$ for the 0-ary relation variables, $F_1^0, F_2^0, \ldots$. If we need 0-ary relation constants, we use $p_1, q_2, \ldots$ instead of $P_1^0, P_2^0, \ldots$.

(.3) We omit parentheses in formulas whenever we possibly can, i.e., whenever we can do so without ambiguity. Thus, we almost always drop outer parentheses and assume $\neg$, $\forall$, $\imath$, $\Box$, and $\mathcal{A}$ apply to as little as possible. Also, we assume that $\rightarrow$ dominates $\neg$. These conventions yield the following examples:

- $\neg Pa \rightarrow bQ$ should be parsed as $((\neg Pa) \rightarrow bQ)$, not $\neg(Pa \rightarrow bQ)$.
- $\forall x Px \rightarrow Qx$ should be parsed as $(\forall x Px) \rightarrow Qx$, not $\forall x(Px \rightarrow Qx)$.
- $\Box Pa \rightarrow bQ$ should be parsed as $(\Box Pa) \rightarrow bQ$, not $\Box(Pa \rightarrow bQ)$
- $\mathcal{A}aP \rightarrow Qb$ should be parsed as $(\mathcal{A}aP) \rightarrow Qb$, not $\mathcal{A}(aP \rightarrow Qb)$.

(.4) We sometimes add parentheses and square brackets to assist in parsing certain formulas and terms.

(**6**) **Metadefinition:** Subformulas. Where φ is any formula, we define a *subformula of* φ recursively as follows:

(.1) φ is a subformula of φ.

(.2) If $[\lambda\, \psi]$, $\neg\psi$, $\forall\alpha\psi$, $\Box\psi$, or $\mathcal{A}\psi$ is a subformula of φ, then ψ is a subformula of φ.

(.3) If $\psi \rightarrow \chi$ is a subformula of φ, then ψ is a subformula of φ and χ is a subformula of φ.[76]

(.4) Nothing else is a subformula of φ.

Given this definition, we may say that ψ is a *proper* subformula of φ just in case ψ is a subformula of φ but not identical to φ. The above definition of *subformula of* has an important consequence:

[76] We later extend our language with definitions that introduce formulas of the form $\psi \& \chi$, $\psi \vee \chi$, and $\psi \equiv \chi$. When these formulas are subformulas of φ, so are ψ and χ. See fact (a) in the final paragraph of (18) below.

Metatheorem ⟨7.1⟩
If ψ is a subformula of χ and χ is a subformula of φ, then ψ is a subformula of φ.

There is a proof in the appendix to this chapter.

It is important to observe that, on this definition, the formula φ and its subformulas are *not* a subformulas of either the term $\imath\nu\varphi$ or the term $[\lambda\nu_1\ldots\nu_n\,\varphi]$ $(n \geq 1)$. φ is a *subformula of* τ is not defined for any term τ other than 0-ary relation terms.[77] Thus, if $\imath\nu\varphi$ or $[\lambda\nu_1\ldots\nu_n\,\varphi]$ occur within a formula ψ, those occurrences of their matrix φ (and its subformulas) are not subformulas of ψ, though other occurrences φ in ψ may be subformulas of ψ. Thus, the formulas $F\imath x\varphi$ and $[\lambda x\,\varphi]b$ *have no proper subformulas*!

(7) **Metadefinition:** A Definition of Subterms and Primary Terms. By contrast, however, since formulas are terms, φ is a *subterm* of the terms $\imath\nu\varphi$ and $[\lambda\nu_1\ldots\nu_n\,\varphi]$. And if $\imath\nu\varphi$ or $[\lambda\nu_1\ldots\nu_n\,\varphi]$ occur somewhere within a formula ψ, those occurrences of φ and its subformulas are subterms of ψ.

The notion of *subterm* will be deployed in various ways in what follows and is an important notion of abstract syntax. To define it, we again use both τ and σ are metavariables ranging over terms, and note that formulas are 0-ary relation terms and so fall in the range of these metavariables. We then define τ is a *subterm of* σ as follows:

(.1) If τ is σ, then τ is a subterm of σ, i.e., every term τ is a subterm of itself.

(.2) If τ is subterm of any of $\kappa_1,\ldots,\kappa_n$, or Π^n $(n \geq 1)$, and σ is either the formula $\Pi^n\kappa_1\ldots\kappa_n$ or the formula $\kappa_1\ldots\kappa_n\Pi^n$, then τ is a subterm of σ.

(.3) If τ is a subterm of the formula φ, and σ is any of the formulas $\neg\varphi$, $\forall\alpha\varphi$, $\Box\varphi$, or $\mathcal{A}\varphi$, then τ is a subterm of σ.

(.4) If τ is a subterm of the formula φ or the formula ψ, and σ is the formula, $\varphi \to \psi$, then τ is a subterm of σ.

(.5) If τ is a subterm of the formula φ, and σ is a λ-expression of the form $[\lambda\nu_1\ldots\nu_n\,\varphi]$ $(n \geq 0)$, then τ is a subterm of σ.

(.6) If τ is a subterm of the formula φ, and σ is a description of the form $\imath\nu\varphi$, then τ is a subterm of σ

Given this definition, we may say that τ is a *proper* subterm of term σ just in case τ is a subterm of σ but not identical to σ. Thus, the terms $\imath x(p \to Gx)$ and $[\lambda x\, p \to Gx]$ have the same *proper* subterms, namely, p, G, x, Gx, and

[77] Complex, 0-ary relation terms of the form $[\lambda\,\varphi]$ do have subformulas. By (6.1), $[\lambda\,\varphi]$ is a subformula of itself. So by (6.2), φ is a subformula of $[\lambda\,\varphi]$. And any subformulas of φ thereby become subformulas of $[\lambda\,\varphi]$.

$p \to Gx$. Clearly, constants and variables are the only terms that have no proper subterms.

Using the definition of *subterm*, we could say:

(.7) A formula φ *contains* a term τ just in case τ is a subterm of φ.

So, for example, the formula $Ry\iota xQx$ contains the terms $Ry\iota xQx$, R, y, ιxQx, Qx, Q, and x.[78] However, $Ry\iota xQx$ contains no λ-expressions, since no λ-expression is a subterm of $Ry\iota xQx$. Moreover, the formula $[\lambda x\, Rxa]y$ contains the terms $[\lambda x\, Rxa]y$, $[\lambda x\, Rxa]$, Rxa, R, x, a, and y (exercise). Neither the λ-expression $[\lambda x\, Rxa]$ nor its matrix Rxa contains a description.

It is also helpful, on occasion, to refer to the *primary terms* of n-ary exemplification and encoding formulas ($n \geq 1$) as follows:

(.8) When $n \geq 1$ and φ is an exemplification formula of the form $\Pi^n\kappa_1 \ldots \kappa_n$ or an encoding formula of the form $\kappa_1 \ldots \kappa_n \Pi^n$, then these occurrences of Π^n, $\kappa_1, \ldots,$ and κ_n are the *primary terms* of (or in) φ.

Thus, the primary terms of $Ra\iota xQx$ are R, a, and ιxQx, and the primary terms of $ab[\lambda xy\, \neg Rxy]$ are a, b, and $[\lambda xy\, \neg Rxy]$. At present, we have defined primary terms only for exemplification and encoding formulas. If φ is such a formula and τ is a primary term of φ, then none of the proper subterms of τ are primary terms of φ.

(**8**) **Metadefinitions:** Operator Scope and Free (Bound) Occurrences of Variables. Let us use β as an additional metavariable ranging over any simple object-language variable. We then define the *scope* of an occurrence of the formula- and term-building operators as follows:

(.1) The formulas $\neg\psi$, $\Box\psi$, and $\mathcal{A}\psi$ are, respectively, the scope of the occurrence of the operators $\neg$, $\Box$ and $\mathcal{A}$. The formula $\psi \to \chi$ is the scope of the occurrence of the operator $\to$.

(.2) The formula $\forall\beta\psi$ is the scope of the left-most occurrence of the operator $\forall\beta$ in that formula. (We say that the matrix ψ is the *proper* scope of $\forall\beta$ in $\forall\beta\psi$.)

(.3) The term $\iota\nu\psi$ is the scope of the left-most occurrence of the operator $\iota\nu$ in that term. (We say that the matrix ψ is the *proper* scope of $\iota\nu$ in $\iota\nu\psi$.)

[78] This can be established as follows: (A) Since $Ry\iota xQx$ is a 0-ary term, it follows by (.1) that $Ry\iota xQx$ is a subterm of itself. So it contains itself. (B) By (.1), R, y, and ιxQx are all subterms of themselves and so by (.2), they are all subterms of $Ry\iota xQx$. So $Ry\iota xQx$ also contains R, y, and ιxQx. (C) By (.1), Qx is a subterm of itself, and so by (.6), it is a subterm of ιxQx. So by (.2), Qx is also a subterm of $Ry\iota xQx$. Hence $Ry\iota xQx$ contains Qx. (D) By (.1), both Q and x are subterms of themselves, and so by (.2), they are subterms of Qx. By (.6), therefore, Q and x are subterms of ιxQx. So by (.2) they are subterms of $Ry\iota xQx$. Hence, $Ry\iota xQx$ contains the terms Q and x.

(.4) The term $[\lambda\nu_1 \ldots \nu_n\, \psi]$ $(n \geq 0)$ is the scope of the left-most occurrence of the operator $\lambda\nu_1 \ldots \nu_n$ in that term. (We say that the matrix ψ is the *proper* scope of $\lambda\nu_1 \ldots \nu_n$ in $[\lambda\nu_1 \ldots \nu_n\, \psi]$.)

Now to define *free* and *bound* occurrences of a variable within a formula or term, we first say that $\forall\beta$ is a variable-binding operator *for* β, that $\imath\nu$ is a variable-binding operator *for* ν, and that $\lambda\nu_1 \ldots \nu_n$ is a variable-binding operator *for* $\nu_1, \ldots, \nu_n$. We then say, for any variable α and any formula φ (or any complex term τ):

(.5) An occurrence of α in φ (or τ) within the scope of an occurrence of a variable-binding operator for α is *bound*; otherwise, the occurrence is *free*.

Finally, we say:

(.6) Those occurrences of β that are free in ψ are *bound by* the left-most occurrence of $\forall\beta$ in $\forall\beta\psi$, as is the occurrence of β in that occurrence of $\forall\beta$; those occurrences of ν that are free in ψ are *bound by* the left-most occurrence of $\imath\nu$ in $\imath\nu\psi$, as is the occurrence of ν in that occurrence of $\imath\nu$; and those occurrences of ν_i $(1 \leq i \leq n)$ that are free in ψ are *bound by* the left-most occurrence of $\lambda\nu_1 \ldots \nu_n$ in $[\lambda\nu_1 \ldots \nu_n\, \psi]$, as are the occurrences of $\nu_1, \ldots, \nu_n$ in that occurrence of $\lambda\nu_1 \ldots \nu_n$.

We henceforth say that:

(.7) A variable α *occurs free* or *is free* in formula φ or term τ if and only if at least one occurrence of α in φ or τ is free, i.e., if and only if φ or τ contains at least one free occurrence of α.

(**9**) **Metadefinitions:** Encoding Position and Core λ-Expressions. To delineate a distinguished group of λ-expressions that play a central role in one of the key axioms of object theory, we first define:

(.1) Term τ *occurs in encoding position in* φ iff τ occurs as a primary term of an encoding formula subterm of φ.

In other words, τ occurs in encoding position in φ just in case φ contains (7.7) an encoding formula the form $\kappa_1 \ldots \kappa_n \Pi^n$ and τ is one of $\kappa_1, \ldots, \kappa_n$, or Π^n. We then say:

(.2) $[\lambda\nu_1 \ldots \nu_n\, \varphi]$ is a *core* λ-expression if and only if no variable bound by the λ occurs in encoding position in φ.

The intuitive idea is that a core λ-expression is one in which encoding formulas are incidental, rather than integral, to the complex exemplification condition formulated by the expression. Axiom (39.2) will stipulate, among other things, that core λ-expressions are significant (i.e., have a denotation). The reader should confirm that in the following, the λ binds a variable in encoding position and so fail to be examples of core λ-expressions:

- $[\lambda x\, \neg(xG \rightarrow p)]$
- $[\lambda x\, aG \rightarrow P\imath z(Rzb \rightarrow xF)]$

The reader should also confirm the following rules of thumb:

- If $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ contains no encoding formulas, then it is a core λ-expression. Examples include $[\lambda x\, \Diamond E!x]$, $[\lambda y\, \neg E!y]$, $[\lambda z\, \forall x Gzx]$, $[\lambda y\, E!y \rightarrow E!y]$, etc.
- Expressions of the form $[\lambda\, \varphi]$ (in which the λ doesn't bind any variables) and of the form $[\lambda \nu_1 \ldots \nu_n \varphi]$ in which none of the variables $\nu_1, \ldots, \nu_n$ occur free in φ are core λ-expressions.
- If φ in $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ contains encoding formulas with free variables that aren't bound by the λ, then $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ is a core λ-expression. Examples include $[\lambda x\, yF]$, $[\lambda x\, yF \rightarrow \neg Fx]$, $[\lambda z\, \forall G(\neg xG \rightarrow \Box Gz)]$, etc.

However, here are three, somewhat more subtle examples of core λ-expressions:

- $[\lambda x\, y[\lambda z\, Rxz]]$ — the x bound by the λ is not in encoding position even though it occurs within the context of the encoding formula $y[\lambda z\, Rxz]$.
- $[\lambda x\, \imath y(Qyx)P]$ — again, the x bound by the λ is not in encoding position even though it occurs within the content of the encoding formula $\imath y(Qyx)P$.
- $[\lambda x\, \Box\forall z(z[\lambda y Fyx] \equiv z[\lambda y Gyx])]$ — the x bound by the λ is not in encoding position anywhere within the matrix.

Later, when we enrich our language by defining new formulas containing &, $\vee$, $\equiv$, $\exists$, $\Diamond$, $=$, etc., there will be more interesting examples of core λ-expressions.

(**10**) **Metadefinitions:** Open/Closed Formulas/Terms.

(.1) A formula φ is *closed* if no variable occurs free in φ; otherwise φ is *open*.

(.2) A formula φ is a *sentence* if and only if φ is a closed formula.

(.3) A term τ is *closed* if no variable occurs free in τ; otherwise, τ is *open*.

Since it is obvious that constants are closed terms and variables are open terms, we typically use 'open' and 'closed' with respect to formulas and other complex terms. If R is a 2-ary relation constant, F a 2-ary relation variable, a an individual constant, and y an individual variable, then $\imath xRxy$ and $\imath xFxa$ are complex individual terms that are open, while $\imath xRxa$ is a complex individual term that is closed. The λ-expressions $[\lambda x\, Rxy]$ and $[\lambda x\, Fxz]$ are complex unary relation terms that are open, while $[\lambda x\, Rxa]$ is a complex, unary relation term that is closed. Furthermore, if P is a unary relation constant, then $P\imath xRxy$ and $P\imath xFxa$ are open formulas, while $P\imath xRxa$ is a closed formula; and $a[\lambda x\, Fxz]$ is an open formula, while $a[\lambda x\, Rxa]$ is closed.

It sometimes helps to think of open complex terms that aren't formulas, such as $\imath xRxy$ and $[\lambda x\, Rxy]$, as *functional terms* whose value is, intuitively, a function of the value of the variable y.[79] Alternatively we may regard such open complex terms as a kind of complex variable; for each choice of y, $\imath xRxy$ may denote a different individual and $[\lambda x\, Rxy]$ may denote a different unary relation. To forestall ambiguity, we always use *variable* in what follows to refer to the *simple* variables listed in (1).

(**11**) **Metadefinitions:** Closures. We say that:

(.1) φ is a *universal closure* or *universal generalization* of ψ whenever φ is the result of prefacing any a string of zero of more universal quantifiers to ψ, i.e., if and only if for some variables $\alpha_1,\ldots,\alpha_n$ $(n \geq 0)$, φ is $\forall\alpha_1 \ldots \forall\alpha_n\psi$.

(.2) φ is an *actualization* of ψ whenever φ is the result of prefacing a string of zero or more occurrences of the actuality operator $\mathcal{A}$ to ψ.

(.3) φ is a *necessitation* or *modal closure* of ψ whenever φ is the result of prefacing any a string of zero or more occurrences of the necessity operator $\Box$ to ψ.

Furthermore, we say:

(.4) φ is a *closure* of ψ if and only if for some variables $\alpha_1,\ldots,\alpha_n$ $(n \geq 0)$, φ is the result of prefacing any string of zero or more occurrences of universal quantifiers, $\mathcal{A}$ operators, and $\Box$ operators to ψ.

Finally, we say:

(.5) φ is a $\Box$-*free closure* of ψ if and only if φ is the result of prefacing any string of zero or more occurrences of actuality operators and universal quantifiers to ψ.

[79] We have taken care to call these expressions *functional terms* instead of *function terms*. The reason is that, strictly speaking, they don't denote functions. A function is, strictly speaking, a binary *relation* R such that $\forall x \exists y(Rxy \,\&\, \forall z(Rxz \rightarrow z = y))$. The expression $[\lambda x\, Rxy]$ doesn't denote such an R. Instead, for each value that y takes, $[\lambda x\, Rxy]$ denotes a unary relation. We therefore refrain from calling $[\lambda x\, Rxy]$ a 'function term'; rather, 'functional term' is more appropriate.

Since we're counting the empty string as a string, the definitions yield that every formula is a closure of itself.

(**12**) **Metadefinition:** Terms of the Same Type. We say that τ and σ are terms *of the same type* iff τ and σ are both individual terms or are, for some $n \geq 0$, both n-ary relation terms.

(**13**) **Remark:** The Symbol '=' in Both Object Language and Metalanguage. Though some texts use different symbols for object-theoretic and metatheoretic identity, we shall use the symbol = for both. Object-theoretic identity is not taken as a primitive; we instead define this notion by cases in item (23) below. By contrast, we formulate metadefinitions in which the symbol '=' represents a primitive *metatheoretic* notion of identity. Metatheoretic identity helps us to describe facts about our system, as opposed to asserting facts within our system. We've already used it in Metadefinition (6) to define *proper* subformulas and in Metadefinition (7) to define *proper* subterms, though in these definitions, we didn't use the formal symbol '='. But now, we'll now use '=' metatheoretically in (14) below to help us define φ^{τ}_{α}, i.e., the result of substituting a term τ for all the free occurrences of the variable α in φ. The syntactic notion φ^{τ}_{α} is subsequently used to state various principles (i.e., definitions, axioms, rules of inference, and theorems).

Once object-theoretic identity is defined, it should always be clear when = is being used object-theoretically to assert something in the object language and when it is being used metatheoretically to assert or describe something in the metalanguage. Of course, if our philosophical project succeeds, the theory of identity developed in the object language ultimately provides an analysis of the primitive metatheoretic notion of identity.

(**14**) **Metadefinition:** Substitutions. The definitions in this item and the next are required to state, for example, the axioms of quantification in item (39), the axioms of identity in item (41), and the axioms for complex relation terms in item (48). In what follows, we use σ as an additional metavariable ranging over any term whatsoever.

- Where α is any variable and τ is any term of the same type as α, we use the notation φ^{τ}_{α} and σ^{τ}_{α}, respectively, to stand for the result of substituting the term τ for every free occurrence of the variable α in formula φ and in term σ.

This notion may be defined more precisely by recursion, based on the syntactic complexity of σ and φ as follows, where the parentheses serve only to eliminate ambiguity and we suppress the obvious superscript indicating arity on the metalinguistic relation variable Π:

- If σ is the variable α and τ is a term of the same type as α, $\sigma^\tau_\alpha = \tau$.
  If σ is a constant or variable other than α and τ is a term of the same type as α, $\sigma^\tau_\alpha = \sigma$.
- If φ is $\Pi\kappa_1 \ldots \kappa_n$, then $\varphi^\tau_\alpha = \Pi^\tau_\alpha \kappa_{1}{}^\tau_\alpha \ldots \kappa_{n}{}^\tau_\alpha$.
  If φ is $\kappa_1 \ldots \kappa_n \Pi$, then $\varphi^\tau_\alpha = \kappa_{1}{}^\tau_\alpha \ldots \kappa_{n}{}^\tau_\alpha \Pi^\tau_\alpha$.
- If φ is $\neg\psi$, $\Box\psi$ or $\mathcal{A}\psi$, then $\varphi^\tau_\alpha = \neg(\psi^\tau_\alpha)$ or $\Box(\psi^\tau_\alpha)$, or $\mathcal{A}(\psi^\tau_\alpha)$, respectively. If φ is $\psi \to \chi$, then $\varphi^\tau_\alpha = \psi^\tau_\alpha \to \chi^\tau_\alpha$.
- If φ is $\forall\beta\psi$, then $\varphi^\tau_\alpha = \begin{cases} \forall\beta\psi, \text{ if } \alpha = \beta \\ \forall\beta(\psi^\tau_\alpha), \text{ if } \alpha \neq \beta \end{cases}$
- If σ is $\imath\nu\psi$, then $\sigma^\tau_\alpha = \begin{cases} \imath\nu\psi, \text{ if } \alpha = \nu \\ \imath\nu(\psi^\tau_\alpha), \text{ if } \alpha \neq \nu \end{cases}$
- If σ is $[\lambda\nu_1 \ldots \nu_n \psi]$, then $\sigma^\tau_\alpha = \begin{cases} [\lambda\nu_1 \ldots \nu_n\, \psi], \text{ if } \alpha \text{ is one of } \nu_1, \ldots, \nu_n \\ [\lambda\nu_1 \ldots \nu_n\, \psi^\tau_\alpha], \text{ if } \alpha \text{ is none of } \nu_1, \ldots, \nu_n \end{cases}$

Note both that (a) σ^τ_α is not defined if α and τ are terms of a different type, and (b) if α doesn't occur free in σ, then $\sigma^\tau_\alpha = \sigma$, and if α doesn't occur free in φ, then $\varphi^\tau_\alpha = \varphi$. The reader should also verify that:

> **Metatheorems** ⟨7.2⟩: $\sigma^\alpha_\alpha = \sigma$ and $\varphi^\alpha_\alpha = \varphi$.

We shall also want to define multiple simultaneous substitutions of terms for variables in φ and σ, but since such a recursive definition would be extremely difficult to read, we simply rest with the following definition: where $\alpha_1, \ldots, \alpha_m$ are any distinct variables and $\tau_1, \ldots, \tau_m$ are any terms of the same types, respectively, as $\alpha_1, \ldots, \alpha_m$, we let $\varphi^{\tau_1,\ldots,\tau_m}_{\alpha_1,\ldots,\alpha_m}$ stand for the result of simultaneously substituting the term τ_i for each free occurrence of the corresponding variable α_i in φ, for each i such that $1 \leq i \leq m$. In other words, $\varphi^{\tau_1,\ldots,\tau_m}_{\alpha_1,\ldots,\alpha_m}$ is the result of making all of the following substitutions simultaneously: (a) substituting τ_1 for every free occurrence of α_1 in φ, (b) substituting τ_2 for every free occurrence of α_2 in φ, etc. Similarly, where $\tau_1, \ldots, \tau_m$ are any terms and $\alpha_1, \ldots, \alpha_m$ are any distinct variables, we let $\sigma^{\tau_1,\ldots,\tau_m}_{\alpha_1,\ldots,\alpha_m}$ stand for the result of simultaneously substituting the term τ_i for each free occurrence of the corresponding variable α_i in σ, for each i such that $1 \leq i \leq m$.

If, for any i, τ_i does not have the same type as α_i, then $\varphi^{\tau_1,\ldots,\tau_m}_{\alpha_1,\ldots,\alpha_m}$ and $\sigma^{\tau_1,\ldots,\tau_m}_{\alpha_1,\ldots,\alpha_m}$ are not well-formed. The reader should verify that:

> **Metatheorems** ⟨7.3⟩: $\sigma^{\alpha_1,\ldots,\alpha_n}_{\alpha_1,\ldots,\alpha_n} = \sigma$ and $\varphi^{\alpha_1,\ldots,\alpha_n}_{\alpha_1,\ldots,\alpha_n} = \varphi$.

(**15**) **Metadefinitions:** Substitutable at an Occurrence and Substitutable For. We say:

- Term τ is *substitutable at an occurrence* of α in formula φ or term σ just in case (a) τ is a term of the same type as α, and (b) that occurrence of α does not appear within the scope of any operator binding a variable that has a free occurrence in τ.

In other words, τ is substitutable at an occurrence of α in φ or σ just in case every occurrence of any variable β free in τ remains an occurrence that is free when τ is substituted for that occurrence of α in φ or σ. Then we say:

- τ is *substitutable for* α in φ or σ just in case τ is substitutable at every free occurrence of α in φ or σ.

In other words, τ is substitutable for α in φ or σ just in case every occurrence of any variable β free in τ remains an occurrence that is free in φ^τ_α or σ^τ_α, i.e., remains an occurrence that is free when τ is substituted for every free occurrence of α in φ or σ.

The following are consequences of this definition:

Metatheorems ⟨7.4⟩:

- Every term τ is trivially substitutable for α in φ if there are no free occurrences of α in φ.
- α is substitutable for α in φ or σ.
- If τ contains no free variables, then τ is substitutable for any variable α in any formula φ or complex term σ.
- If none of the free variables in τ occur bound in φ or σ, then τ is substitutable for any α in φ or σ.

Note: The prime symbol $'$ is used to avoid overspecificity. When we attach a prime symbol to a metavariable, the resulting metavariable stands for an object-language expression that may be distinct from, and not necessarily related to, the expression signified by the metavariable without the prime. In the next item, we shall use φ' to stand for an alphabetical-variant of the formula φ, and use τ' to stand for an alphabetical-variant of the term τ. In the next chapter, the axiom for the substitution of identicals (41) will use φ' to indicate the result of replacing zero or more free occurrences of the variable α in φ with occurrences of the variable β. (So in the case where zero occurrences are replaced, φ' is φ.) Later we shall use the metavariable ρ' to denote any η-variant of the relation term ρ. Sometimes we shall place primes on expressions in the object language; for example, in a later chapter, c is introduced as a restricted variable in the object-language that ranges over classes, and we let $c', c'', \ldots$ be *distinct* restricted variables for classes (and so on for other restricted variables). The context should always make it clear how the prime symbol $'$ is being used.

(**16**) **Metadefinitions:** Alphabetic Variants. For basic examples of *alphabetic variants*, some readers may wish to check the opening of explanatory Remark (26). The following definition presupposes an intuitive understanding of the concept we're trying define.[80]

To precisely define the general notion of *alphabetic variant*, i.e., for formulas and terms of arbitrary complexity, we first define *linked* and *independent* occurrences of a variable.

(.1) Let α_1 and α_2 be occurrences of the variable α in the formula φ or in term τ. Then we say that α_1 is *linked to* α_2 in φ or τ (or say that α_1 and α_2 are *linked* in φ or τ) just in case:

(a) either both α_1 and α_2 are free, or

(b) both α_1 and α_2 are bound by the same occurrence of a variable-binding operator.

Otherwise, we say that α_1 and α_2 are *independent* in φ or τ.

Examples of *linked* and *independent* variables are given in Remark (26). (.1) gives rise to:

Metatheorem ⟨7.5⟩
Linked is an equivalence condition on variable occurrences in a formula (or term).

(A proof is given in the Appendix to this chapter.) Thus, the occurrences of each variable in a formula (or term) can be partitioned into *linkage groups*. Each linkage group is a cell of the partition; in a linkage group for a variable α in φ or τ, each occurrence of α in the group is linked to every other occurrence,

[80] Traditionally, two formulas or complex terms are defined to be *alphabetic variants* just in case some sequence of uniform permutations of the bound variables (in which no variable is captured during a permutation) transforms one expression into the other. So, for example, $\forall F(Fx \equiv Fy)$ and $\forall G(Gx \equiv Gy)$ would be alphabetically-variant formulas because the permutation sequence $F \to G$ transforms the first formula into the second. Similarly, $\imath x\forall yMyx$ and $\imath y\forall zMzy$ would be alphabetically-variant terms because the permutation sequence $y \to z, x \to y$ transforms the first description into the second. And $[\lambda yRy\imath zQz]$ and $[\lambda zRz\imath xQx]$ are variant λ-expressions because the permutation sequence $z \to x, y \to z$ transforms the first λ-expression into the second. However, we shall *not* follow the traditional definition but rather develop a definition that identifies alphabetic variants by way of symmetries among bound variable occurrences.

while occurrences of α in different linkage groups are independent of one another.

We now introduce some notation ('BV-notation') that makes explicit the occurrences of *bound* variables in formulas and terms:

(.2) When $\alpha_1, \ldots, \alpha_n$ is the list of all the variable occurrences bound in formula φ or complex term τ, in order of occurrence and including repetitions of the same variable, the BV-notation for φ is $\varphi[\alpha_1, \ldots, \alpha_n]$, and the BV-notation for τ is $\tau[\alpha_1, \ldots, \alpha_n]$, respectively, .

The reader is encouraged to examine the examples of BV-notation given in Remark (26).

Next we introduce notation for replacing bound variables:

(.3) We write $\varphi[\beta_1/\alpha_1, \ldots, \beta_n/\alpha_n]$ to refer to the result of replacing α_i by β_i in $\varphi[\alpha_1, \ldots, \alpha_n]$, for $1 \leq i \leq n$. Analogously, we write $\tau[\beta_1/\alpha_1, \ldots, \beta_n/\alpha_n]$ to refer to the result of replacing α_i by β_i in term $\tau[\alpha_1, \ldots, \alpha_n]$.

Finally, we may define:[81]

(.4) φ' is an *alphabetic variant* of φ just in case, for some n:

- $\varphi' = \varphi[\beta_1/\alpha_1, \ldots, \beta_n/\alpha_n]$,
- φ' has the same number of bound variable occurrences as φ and so can be written as $\varphi'[\beta_1, \ldots, \beta_n]$, and
- for $1 \leq i, j \leq n$, α_i and α_j are linked in $\varphi[\alpha_1, \ldots, \alpha_n]$ if and only if β_i and β_j are linked in $\varphi'[\beta_1, \ldots, \beta_n]$

(.5) τ' is an *alphabetic variant* of τ just in case, for some n:

- $\tau' = \tau[\beta_1/\alpha_1, \ldots, \beta_n/\alpha_n]$,
- τ' has the same number of bound variable occurrences as τ and so can be written as $\tau'[\beta_1, \ldots, \beta_n]$, and
- for $1 \leq i, j \leq n$, α_i and α_j are linked in $\tau[\alpha_1, \ldots, \alpha_n]$ if and only if β_i and β_j are linked in $\tau'[\beta_1, \ldots, \beta_n]$.

Additional examples of alphabetic variants that illustrate the above definitions are given in Remark (26).

Note that our definitions require that if β is to replace α to produce an alphabetic variant of a formula φ (or term τ), then β must not occur free within the scope of a variable-binding operator that binds an occurrence of α in φ (or

[81] I'd like to thank Daniel West for noting that a problem with an earlier version of (.4) and (.5) – something like the second condition was missing. And thanks to Uri Nodelman for his suggestion that the second condition is probably the simplest solution to the problem.

τ). For example, in the formula $\forall xRxy$ (= φ), y occurs free within the scope of $\forall x$. We don't obtain an alphabetic variant of φ by substituting occurrences of y for the occurrences of x in the linkage group of bound occurrences of x. The formula that results from such a replacement, $\forall yRyy$ (= φ'), is very different in meaning from the original. In this example, the single occurrence of y in φ gets captured when occurrences of y replace the bound occurrences of x. Thus, φ in BV-notation is $\varphi[x,x]$ and φ' in BV-notation is $\varphi'[y,y,y]$, and so the first condition in the definition of *alphabetic variant* fails because the number of bound variables, counting multiple occurrences, is different for φ and φ'. There are a number of observations about alphabetic variants at the end of Remark (26) that might prove helpful.

7.2 Definitions Extending the Object Language

In what follows, we regularly state definitions that extend our object language. These introduce new expressions into our language. Some of the axioms and axiom schemas of object theory will be stated in terms of these new expressions. In a definition extending the object language, a *new* expression, the *definiendum*, is introduced by way of a *definiens* that contains only primitive expressions or previously defined expressions. We shall not regard definitions as metalinguistic abbreviations of the object language but rather as conventions for (a) extending the object language with new syncategorematic expressions, formulas, and terms, and (b) conservatively extending our deductive system with new and *safe* axioms. As such, theorems stated in terms of defined notions become genuine philosophical statements of the object language rather than metaphilosophical statements of the metalanguage. In this section, we restrict our attention to how definitions achieve (a), though on occasion, we'll have to make reference to how they achieve (b). This latter is discussed in Section 9.4, where we carefully characterize the inferential role that such definitions play within our deductive system.

(**17**) **Remark:** Definitions by Equivalence and Definitions by Identity. In a system such as the present one, in which identity is defined rather than primitive and in which there are both individual and relation terms that may fail to have a denotation, definitions have to be formulated carefully so as to avoid misunderstanding.[82] So it is important to begin by laying out at least some of the

[82] Though Frege's insightful discussion of definitions informed his second-order system, most discussions of the theory of formal definitions have been framed with respect to the language of the first-order predicate calculus with identity, sometimes extended by function terms. In thinking about the theory of definitions, I consulted Frege 1879, §24; Padoa 1900; Frege 1903a, §§55–67, §§139–144, and §§146–147; Frege 1903b, Part I; Frege 1914, 224–225; Suppes 1957; Mates 1972; Dudman 1973; Belnap 1993; Hodges 2008; Urbaniak & Hämäri 2012; and Gupta 2014. Hodges

requirements for formulating definitions in the object language and at least some of the conventions that govern them. We'll postpone discussion of the inferential role of definitions until Chapter 9, in items (72) and (73).

There will be two types of definition in our system: Definition by Equivalence and Definition by Identity. We'll refer to them as *definitions-by-*$\equiv$ and *definitions-by-*$=$, respectively. As a rough, initial characterization of the distinction, we can say that the former stipulates an equivalence between formulas, whereas the latter stipulates an identity between terms. Thus, a definition-by-$\equiv$ has the form:

$$\varphi \equiv_{df} \psi$$

In the general case, where there are n-free variables $\alpha_1, \ldots, \alpha_n$ in the definiens and definiendum, a definition-by-$\equiv$ has the form:

$$\varphi(\alpha_1, \ldots, \alpha_n) \equiv_{df} \psi(\alpha_1, \ldots, \alpha_n)$$

By contrast a definition-by-$=$ has the form:

$$\tau =_{df} \sigma$$

provided τ and σ are terms of the same type, neither of which is a variable. In the general case, where there are n-free variables $\alpha_1, \ldots, \alpha_n$ in the definiens and definiendum, a definition-by-$=$ has the form:

$$\tau(\alpha_1, \ldots, \alpha_n) =_{df} \sigma(\alpha_1, \ldots, \alpha_n)$$

provided τ and σ are terms of the same type.

Note that it would be incorrect to *distinguish* definitions-by-$\equiv$ and definitions-by-$=$ by saying that the former introduce new formulas and the latter introduce new terms. This oversimplification is undermined by the fact that all and only formulas are 0-ary relation terms, a fact which has the following consequences. First, the definiendum φ and definiens ψ in a definition-by-$\equiv$ are 0-ary relation terms as well as formulas – so in every case, these definitions introduce equivalences between 0-ary relation terms. Second, the definiendum τ and definiens σ in a definition-by-$=$ may be formulas if τ and σ are 0-ary relation terms – so in some cases, these definitions introduce an identity between formulas.

In light of these observations, we emphasize that the distinction between definitions-by-$\equiv$ and definitions-by-$=$ concerns their inferential role in the deductive system. The inferential role of both kinds of definition has to be carefully formulated and this will be done in Chapter 9, as part of the development of the deductive system. For now, we can describe their role pre-theoretically

2008 and Urbaniak & Hämäri 2012 provide insightful discussions of the contributions by Kotarbiński, Łukasiewicz, Leśniewski, Ajdukiewicz and Tarski to the elementary theory of definitions.

as follows: (a) a definition-by-≡ implicitly introduces necessary biconditionals (and their closures) as axioms,[83] and (b) a definition-by-= implicitly introduces conditional identities (and their closures) as axioms, where the identities are conditioned on the claim that the definiens has a denotation. Since biconditionals and conditional identities have very different inferential roles, the inferences one can draw from the two forms of definition will be very different. Further discussion of this point, however, is premature, for we can't even state the inferential role of definitions-by-= until we define formulas of the form $\tau = \sigma$.

As noted above, both types of definition may contain free variables, and so the first requirement for formulating definitions is that the definiendum and definiens should have matching free variables:

(.1) **Matching Free Variables Requirement:**
In any definition, all and only the variables that occur free in a definiens should also occur free in the definiendum.

For some purposes, it proves useful to relax this requirement by allowing the definiendum to contain free variables that aren't free in the definiens.[84] But there are well-known reasons for not allowing the definiens to contain free variables that aren't free in the definiendum.[85]

Next we adopt a convention that assumes familiarity with the various reasons why one might want to cast definitions as schemata involving metavariables. Readers unfamiliar with those reasons are encouraged to jump ahead to Section 7.3 (Explanatory Remarks) and read Remarks (27) and (28). Remark (27) explains why, in a system containing complex terms that might fail to have a denotation, any free variables in a definition should be cast as metavariables instead of object-language variables. Remark (28) explains why, in the case of bound variables, metavariables should be used (subject to certain restrictions) instead of object-language variables.

However, definition schemata involving free and bound Greek metavariables are often much more difficult to read and grasp. So although definitions should, strictly speaking, be stated as schemata using metavariables, we shall use metavariables in definitions only when it is clearer to do so, i.e., only when

[83]In the first instance, such definitions implicitly introduce the closures of conditionals, and it will be derivable, from the fact that the definition $\varphi \equiv_{df} \psi$ introduces the closures of $\varphi \to \psi$ and $\psi \to \varphi$, that the definition yields $\Box(\varphi \equiv \psi)$ and its closures as theorems.

[84]Suppes (1957) explains why one can allow definitions in which variables occur free in the definiendum but not in the definiens; one can trivially get the variables to match by adding dummy clauses to the definiens. For example, Suppes notes (1957, 157) that the number-theoretic definition $Q(x,y) =_{df} x > 0$ can be turned into $Q(x,y) =_{df} x > 0 \;\&\; y = y$.

[85]Suppes (1957) has a nice discussion of why the definiens may not contain a free variable that is not free in the definiendum. See his example (1957, 157) of how to derive a falsehood from a definition such as $R(x) \equiv_{df} x + y = 0$.

doing so lightens cognitive load. For the cases where it doesn't, we state and adopt the following convention that will help mitigate the problem and make definitions in the object language much more readable and user-friendly:

(.2) **Convention for Variables in Definitions:**
When it increases clarity and faciliates readability, we formulate definitions with both free and bound object-language variables that function as metavariables, with the understanding that they give rise to *instances* of the definition as follows:

(.a) if object-language variables $\alpha_1,\ldots,\alpha_n$ ($n \geq 0$) occur free in a definition of the form $\varphi \equiv_{df} \psi$ or $\tau =_{df} \sigma$, then where $\tau_1,\ldots,\tau_n$ are *any* terms substitutable, respectively, for $\alpha_1,\ldots,\alpha_n$ in ψ or σ:

$$\varphi^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n} \equiv_{df} \psi^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$$

is an instance of a definition-by-$\equiv$ that extends the language with new formulas of the form $\varphi^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$, and

$$\tau^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n} =_{df} \sigma^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$$

is an instance of a definition-by-$=$ that extends the language with new terms of the form $\tau^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$,[86] and

(.b) if any object-language variables occur bound in the definiens of an instance of a definition of the form $\varphi \equiv_{df} \psi$ or $\tau =_{df} \sigma$, then the result of replacing the definiens of that instance by one of its alphabetic variants is also an instance of the definition.[87]

Thus, whenever the use of metavariables would make a definition significantly more difficult to read and grasp, we may use object-language variables instead. The two clauses of (.2) will become clearer when we start formulating definitions that contain free or bound object-language variables; in the first examples of such definitions, we include, or point to, explanatory remarks that outline how the two clauses of (.2) operate.

[86] In more technical terms, the import of (.a) is that each distinct object-language variable α that occurs free in a definition functions as a distinct Greek metavariable that ranges over terms of the same type as α, with the proviso that if α falls within the scope of a variable-binding operator **Op** in the definiens, then a term may serve as an instance of α only if it doesn't contain free occurrences of the variable bound by **Op**. We shall illustrate this technical description in Remark (31), after we've introduced definition (20.1). Note separately that when the propositional variables $p, q, \ldots$ occur free in the definiens and definiendum, they function either as metavariables for 0-ary relation terms $\Pi^0_1, \Pi^0_2, \ldots$ or as metavariables for formulas $\varphi, \psi, \ldots$.

[87] In more technical terms, the import of (.b) is that each distinct object-language variable α that occurs bound by a variable-binding operator **Op** in the definiens functions as a distinct Greek metavariable ranging over variables of the same type as α, with the proviso that an object-language variable may serve as an instance of α only if it does not occur free in any term occurring within the scope of **Op**. We shall illustrate this technical description in Remark (32), after we've introduced definition (20.1).

Finally, we also adopt the following convention, regarding definitions in which the definiens contains free variables that occur in encoding position (9.1):

(.3) **Encoding Formula Convention:** In any definition in which a variable, say α, occurs free, we regard the definiendum as having a free occurrence of α in encoding position whenever α occurs free in encoding position anywhere in the definiens.

It would serve well to give examples of how this convention will be used, even though the discussion of these examples will reference some defined notions we've not yet introduced (including &, $\vee$, $\equiv$, and =). In (23.1), we define x *is identical to* y ($x\!=\!y$) as follows:

$$x\!=\!y \ \equiv_{df} \ (O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)) \vee (A!x \,\&\, A!y \,\&\, \Box\forall F(xF \equiv yF)) \qquad (23.1)$$

This definitions stipulates that $x = y$ if and only if either x and y are both ordinary that necessarily exemplify the same properties or both are abstract that necessarily encode the same properties. Clearly, the variables x and y occur free in the definiens and in the definiendum. But notice also that in the definiens, there are occurrences of x and y in encoding position. So by the Encoding Formula Convention, the occurrences of x and y in the definiendum $x\!=\!y$ are to be regarded as being in encoding position. Thus, the λ-expression $[\lambda xy\, x\!=\!y]$ is not a core λ-expression as defined in (9.2); the λ binds variables that occur in encoding position in the matrix. This λ-expression attempts to build encoding conditions into the exemplification conditions of a relation and this violates an intuition underlying object theory; one may not introduce such relations until it can be established that it is safe to do so. Since $[\lambda xy\, x\!=\!y]$ isn't a core λ-expression, it does not meet the conditions of axiom (39.2) and so that axiom will not assert that $[\lambda xy\, x\!=\!y]$ signifies a relation.[88]

Taking the example a step further, consider the expressions $[\lambda x\, P\imath x(x = y)]$ and $[\lambda y\, P\imath x(x = y)]$, where P is any property. The first of these is a core λ-expression: since no occurrences of x in the matrix are bound by the λ, no variable bound by the λ occurs in encoding position in the matrix. By contrast, $[\lambda y\, P\imath x(x\!=\!y)]$ fails to be a core λ-expression; the λ binds a variable, namely y,

[88]The expression $[\lambda z\, [\lambda y\, Fyx] = [\lambda y Gyx]]$ is a complex and interesting example. The matrix $[\lambda y\, Fyx]\!=\![\lambda y Gyx]$ is defined in (23.2) as $F{\downarrow} \,\&\, G{\downarrow} \,\&\, \Box\forall z(z[\lambda y\, Fyx] \equiv z[\lambda y\, Gyx])$. In this matrix, the variable x does not occur in encoding position, as we saw when we considered a similar example at the end of (9); intuitively, x occurs in exemplification position in Fyz and Gyz. (We need not consider the fact that $F{\downarrow}$ and $G{\downarrow}$ are both defined terms as well – their definition in (20.2) doesn't involve a free occurrence of z.) So, by our Encoding Formula Convention, x does not occur in encoding position in the matrix $[\lambda y\, Fyx]\!=\![\lambda y Gyx]$. Indeed, both λ-expressions flanking the identity sign qualify as core λ-expressions — no variable bound by the λ occurs in encoding position in the matrix. And, given our Encoding Formula Convention, $[\lambda z\, [\lambda y\, Fyx] = [\lambda y Gyx]]$ is a core λ-expression. Axiom (39.2) will assert that these expressions all have a denotation.

that occurs in encoding position in the matrix $P\imath x(x=y)$. For, as we just saw, we must regard y in this last formula as occurring in encoding position, by our Encoding Formula Convention and the definition of $x=y$. Consequently (to anticipate), axiom (39.2) will assert that the first, but not the second, of our two examples denotes a relation.[89]

These issues iterate as we nest definitions. Consider the following definition, which we'll introduce later purely for illustrative purposes in Remark (27):

$$\iota_y =_{df} \imath x(x=y)$$

This defines *the* y as: the x such that x is identical to y. By the Encoding Formula Convention, since the free occurrence of y occurs in encoding position in the definiens $\imath x(x=y)$, we must regard the free occurrence of y as occurring in encoding position in the definiendum ι_y. So $[\lambda y\, P\iota_y]$ is not a core λ-expressions, as defined in (9.2). But $[\lambda x\, P\iota_y]$ us. To anticipate again, axiom (39.2) will assert that $[\lambda x\, P\iota_y]$ signifies a relation, but does not assert that $[\lambda y\, P\iota_y]$ signifies a relation.

One final point concerns the introduction of new constants using definientia that (a) contain encoding formulas as subterms but (b) have no free variables. For example, the following definitions would be acceptable, though hardly useful:

$$a =_{df} \imath x(xP \;\&\; \neg xP)$$

$$Q =_{df} [\lambda x\, \exists F(xF \;\&\; \neg Fx)]$$

These are both legitimate definitions even though the definiens, in each case, provably fails to have a denotation (the matrix of the definiens in the second definition plays a role in the Clark-Boolos paradox). Our theory of definitions will therefore imply that both a and Q also fail to denote.

Whereas the definiens, in both cases, has an encoding formula as a subterm, the occurrences of the variable x in the encoding formula subterms xP and xF are *bound*. So if we now form the expressions $[\lambda x\, Rxa]$ and $[\lambda x\, Qx]$, where a and Q are defined as above, no variable bound by the λ occurs in encoding position in the matrix. $[\lambda x\, Rxa]$ and $[\lambda x\, Qx]$ are core λ-expressions, as defined in (9.2); by axiom (39.2), they denote properties, albeit necessarily unexemplified properties.

---

[89] Intuitively, $[\lambda x\, P\imath x(x=y)]$ signifies either a property that everything exemplifies or a property that nothing exemplifies, depending on the value of y. Since it a theorem that $y=\imath x(x=y)$ (177.2), the λ-expression has the same exemplification conditions as the property: *being an individual x such that y exemplifies P* ($[\lambda x\, Py]$). If the value assigned to y is an individual that exemplifies P, then $[\lambda x\, P\imath x(x=y)]$ denotes a property that every object exemplifies, and if the value assigned to y is an individual that doesn't exemplify P, then the λ-expression denotes a property that no object exemplifies.

In what follows, it will be crucial to keep a strict eye on the status of the individual variables when introducing definienda in terms of definientia that contain free variables in encoding position. This is key to understanding what is asserted by axiom (39.2). See the discussion immediately following the introduction of this axiom and Remark (55) for further discussion.

7.2.1 The Classical, Sentence-Forming Operators

(**18**) **Definitions:** The Connectives, Existential Quantifier, and Possibility Operator of Classical Quantified Modal Logic. Since it is clearer to use metavariables to define the standard connectives, existential quantifier, and possibility operator of classical quantified modal logic, we introduce them as follows: (.1) φ *and* ψ just in case it is not the case that if-φ-then-not-ψ; (.2) φ *or* ψ just in case if not-φ then ψ; (.3) φ *if and only if* ψ just in case both if-φ-then-ψ and if-ψ-then-φ; (.4) *there exists an* α *such that* φ just in case it is not the case that every α is such that not-φ; and (.5) *It is possible that* φ if and only if it is not the case that necessarily not-φ:

(.1) $\varphi \,\&\, \psi \;\equiv_{df}\; \neg(\varphi \rightarrow \neg\psi)$

(.2) $\varphi \vee \psi \;\equiv_{df}\; \neg\varphi \rightarrow \psi$

(.3) $\varphi \equiv \psi \;\equiv_{df}\; (\varphi \rightarrow \psi) \,\&\, (\psi \rightarrow \varphi)$

(.4) $\exists\alpha\varphi \;\equiv_{df}\; \neg\forall\alpha\neg\varphi$

(.5) $\Diamond\varphi \;\equiv_{df}\; \neg\Box\neg\varphi$

When we uniformly substitute formulas of the object language for φ and ψ in the above and uniformly substitute object-language variables for α, we obtain instances of the definitions.

These definitions preserve the classical understanding of the sentence-forming operators $\&$, $\vee$, $\equiv$, $\exists\alpha$, and $\Diamond$.[90] Where where φ and ψ are any formulas, and α is any variable, the above extend our language with new formulas of the form $\varphi \,\&\, \psi$, $\varphi \vee \psi$, $\varphi \equiv \psi$, $\exists\alpha\varphi$, and $\Diamond\varphi$. We henceforth call both $\Box$ and $\Diamond$ *modal* operators.

We now stipulate that the metadefinitions of *subformula*, *subterm*, *primary term*, *scope*, and *free/bound occurrence*, *open/closed* formulas and terms, *closures*, *substitutable for*, *alphabetic variant*, etc., are to be appropriately extended to ensure that there are facts of the following kinds, among many others: (a) φ is a subformula of $\exists\alpha\varphi$ and $\Diamond\varphi$, and φ and ψ are subformulas of $\varphi\&\psi$, $\varphi\vee\psi$, and

[90] Interested readers might wish to review explanatory Remark (29), where we adduce reasons why one could have introduced this definition by using object-language variables under Convention (17.2).

$\varphi \equiv \psi$; (b) if τ is a subterm of φ or ψ, then τ is a subterm of $\varphi \,\&\, \psi$, $\varphi \vee \psi$, $\varphi \equiv \psi$, $\exists\alpha\varphi$, and $\Diamond\varphi$; (c) the formulas $\psi \,\&\, \chi$, $\psi \vee \chi$, and $\psi \equiv \chi$ are, respectively, the scope of the occurrence of the operators &, $\vee$, and $\equiv$; (d) the formula $\exists\beta\psi$ is the scope of the left-most occurrence of the operator $\exists\beta$ in that formula, and ψ is that operator's proper scope; (e) the formula $\Diamond\psi$ is the scope of the occurrence of the operator $\Diamond$ in that formula; and (f) those occurrences of β that are free in ψ are *bound by* the left-most occurrence of $\exists\beta$ in $\exists\beta\psi$, as is the occurrence of β in that occurrence of $\exists\beta$; and so on.

(**19**) **Notational Conventions:** Dominance Order. In what follows, our binary connectives shall be governed by the following partial *dominance order*: $\equiv$ dominates $\rightarrow$, $\rightarrow$ dominates both & and $\vee$, and neither & nor $\vee$ dominate each other. For example:

- $\varphi \rightarrow \psi \equiv \neg\psi \rightarrow \neg\varphi$ should be parsed as $(\varphi \rightarrow \psi) \equiv (\neg\psi \rightarrow \neg\varphi)$
- $\varphi \,\&\, \psi \rightarrow \chi$ should be parsed as $(\varphi \,\&\, \psi) \rightarrow \chi$
  $\varphi \vee \psi \rightarrow \chi$ should be parsed as $(\varphi \vee \psi) \rightarrow \chi$
- $(\varphi \vee \psi) \,\&\, \chi$ and $\varphi \vee (\psi \,\&\, \chi)$ must be written as they are.

In cases such as the first two, we may drop parentheses without ambiguity.

7.2.2 Existence in the Logical Sense

(**20**) **Definitions:** Existence in the Logical Sense. We now define a condition that asserts *existence* in the *logical* sense. This condition is distinct from the variable-binding quantifier $\exists\alpha$ (which asserts 'there exists an α such that') and is distinct from the *physical* sense in which individuals are sometimes said to exist. The quantifier $\exists\alpha$ was defined in (18.4) and the physical sense of existence is already represented in our system by the distinguished term $E!$, which may be interpreted as denoting the primitive property *being concrete*. By contrast, we now define a logical sense in which individuals and relations *exist*. Some readers may find it useful to first read Remark (30), which attempts to describe what is distinctive about the following definitions. The discussion there contrasts our approach, in which existence is defined in terms of predication, with classical approaches in which existence is defined in terms of identity. As we shall see, both logical existence *and* identity are definable in terms of predication in the present system.

We use the notation $\downarrow$ to represent *existence* and define this notion for the following three, mutually-exclusive and jointly-exhaustive cases: (.1) individual terms, (.2) n-ary relation terms ($n \geq 1$), and (.3) 0-ary relation terms. By employing Convention (17.2), we cast these definitions using object-language variables.

We first define (.1) x *exists* just in case x exemplifies some property:

(.1) $x{\downarrow} \equiv_{df} \exists F Fx$

By Convention (17.2), the following are perfectly good instances of this definition:

- $y{\downarrow} \equiv_{df} \exists G Gy$
- $\imath zFz{\downarrow} \equiv_{df} \exists H H\imath zFz$

However, $\imath zFz{\downarrow} \equiv_{df} \exists F F\imath zFz$ is not an instance; the term $\imath zFz$ is not substitutable for x in the definiens of (.1) since its free variable F would get captured by the quantifier $\exists F$. These facts make it clear why (.1), under Convention (17.2), is equivalent to the schematic definition $\kappa{\downarrow} \equiv_{df} \exists\Omega\,\Omega\kappa$, where κ doesn't contain any free occurrences of the unary relation variable Ω. For a full discussion of this point, see Remarks (31) and (32), where we explain, respectively, the substitutability requirement in Convention (17.2.a) and the alphabetically-variant definientia allowed by Convention (17.2.b).

For $n \geq 1$, we say (.2) F^n *exists* just in case there are individuals $x_1,\ldots,x_n$ that encode F^n:

(.2) $F^n{\downarrow} \equiv_{df} \exists x_1 \ldots \exists x_n(x_1 \ldots x_n F^n)$ $\qquad (n \geq 1)$

Given Convention (17.2), the following are perfectly good instances of this definition:

- $G{\downarrow} \equiv_{df} \exists y(yG)$
- $[\lambda x\, Rxy]{\downarrow} \equiv_{df} \exists z(z[\lambda x\, Rxy])$

However, $[\lambda x\, Rxy]{\downarrow} \equiv_{df} \exists y(y[\lambda x\, Rxy])$ is not an instance. The variable y occurs free in $[\lambda x\, Rxy]$ and so one may not use y as the variable bound by the quantifier $\exists$ in the definiens, since that captures the free occurrence of y in the definiendum. In general, Convention (17.2) allows us to produce well-formed instances of this definition by uniformly replacing F^n with any n-ary relation term and by uniformly replacing $x_1,\ldots,x_n$, respectively, with any distinct individual variables $\nu_1,\ldots,\nu_n$, provided none of the ν_i occur free in the term replacing F^n.

Interested readers might wish to skip ahead to explanatory Remark (33), to get some idea as to why (.2) offers a good definition of n-ary relation existence for $n \geq 1$. Moreover, readers interested in why n-ary relation existence for $n \geq 2$ can't be defined in terms of property existence are referred to explanatory Remark (34).

When $n = 0$, we say (.3) (the proposition) p *exists* just in case (the property) *being an individual such that* p exists:

(.3) $p{\downarrow} \equiv_{df} [\lambda x\, p]{\downarrow}$

It is important here that the variable x is vacuously bound by the λ in the definiens. If we consider the 0-ary relation term $P\imath zQz$, then by the clauses of Convention (17.2), the following is an instance of (.3):

$$P\imath zQz{\downarrow} \equiv_{df} [\lambda z\, P\imath zQz]{\downarrow}$$

This is an instance because the λ vacuously binds z – there are no free occurrences of z in the matrix $P\imath zQz$ and so no occurrences of z are bound by λz.[91]

By contrast, the following is not an instance of (.3):

$$Rxy{\downarrow} \equiv_{df} [\lambda x\, Rxy]{\downarrow}$$

Here, the free occurrence of x in Rxy gets captured in the definiens by the variable-binding operator λx. To define $Rxy{\downarrow}$ as specified by (.3), we have to formulate an instance such as $Rxy{\downarrow} \equiv_{df} [\lambda z\, Rxy]{\downarrow}$.

It is important to mention, in connection with (.3), that for some 0-ary relation terms, the definiendum may need disambiguating parentheses. For example, the formula $\neg Px{\downarrow}$ is ambiguous between $\neg(Px{\downarrow})$ and $(\neg Px){\downarrow}$, given that both Px and $\neg Px$ are 0-ary terms. If, in any given case, parentheses don't disambiguate which is meant, then we employ the convention: $\imath x\varphi{\downarrow}$ designates $(\imath x\varphi){\downarrow}$; otherwise $\downarrow$ applies to the smallest unit possible, so that $\neg Px{\downarrow}$ is to be interpreted as $\neg(Px{\downarrow})$.

Finally, now that we have defined $\tau{\downarrow}$ for every term τ, we note that when τ is a term introduced by definition, we sometimes say that τ is *well-defined* if it is a theorem that $\tau{\downarrow}$.

(**21**) **Remark:** An Intentional Ambiguity. In what follows, it shall be useful to have some word in our metalanguage that that is intentionally neutral between the material mode notion of *existence* and the formal mode notion of *having a denotation*, i.e., some studiously ambiguous word that refers both to the material mode claim $\tau{\downarrow}$ ($\ulcorner\tau$ *exists*$\urcorner$) as well as to the formal mode claim that τ has a denotation.[92] We shall use the expressions *significant*, *logically significant*, or *logically proper* for this purpose, and use *empty* for terms that fail to be significant. Given such usage, it will be important to distinguish an empty property term (i.e., a unary relation term τ that doesn't have a denotation) from

[91] Note that the definiendum in this example is not ambiguous, despite the fact that $\imath zQz{\downarrow}$ is well-formed. The expression $P\imath zQz{\downarrow}$ can be interpreted only as $(P\imath zQz){\downarrow}$; it can't be interpreted as $P(\imath zQz{\downarrow})$.

[92] The Quine corner quotes indicate that what τ *exists* asserts is not a metatheoretic claim about the term τ, but a claim in the object language. We shall, in what follows, generally omit the Quine corner quotes around schematic notions involving metavariables, with the understanding that our intention is to define an object-theoretic notion that applies to objects and relations, not a metatheoretic notion that applies to terms. When we do use the Quine corner quotes, this is simply to remind the reader that we are talking about an object-theoretic notion and not a metatheoretic notion.

an empty property (i.e., a property F that isn't exemplified).[93] We'll introduce axioms that guarantee that (.1) only significant terms can be instantiated into universal claims – see axiom (39.1); (.2) the following terms, in the first instance, are stipulated to be significant: primitive constants, variables, and core λ-expressions (9.2), in which no variable bound by the λ occurs in encoding position in the matrix (9.1)– see axiom (39.2); and (.3) (the primary terms of) true exemplification formulas are significant and that (the primary terms of) true encoding formulas are significant – see axioms (39.5.a) and (39.5.b).

7.2.3 Identity

To define *identity* generally, we first need to define the properties of *being ordinary* ($O!$) and *being abstract* ($A!$). These take place in (22.1) and (22.2), respectively. Once these definitions are in place, we may then define, for any terms τ and σ, the condition $\ulcorner\tau$ *is identical to* $\sigma\urcorner$ (written: $\tau = \sigma$), by cases. The four cases of the definition of $\tau = \sigma$ are:

- τ and σ are both individual terms (23.1)
- τ and σ are both unary relation terms (23.2)
- τ and σ are both n-ary relation terms ($n \geq 2$) (23.3)
- τ and σ are both 0-ary relation terms (23.4)

The preliminary definitions of $O!$ and $A!$ are used only in the first of these four cases.

(22) **Definitions:** Ordinary vs Abstract Objects. We may define *being ordinary* ('$O!$') as a new unary relation constant as follows:

(.1) $O! =_{df} [\lambda x\, \Diamond E!x]$

In other words, *being ordinary* is defined as *possibly exemplifying concreteness*. As we shall see, this definition implies $O! = [\lambda x\, \Diamond E!x]$; this will follow from our definitions-by-= and the fact that $[\lambda x\, \Diamond E!x]\downarrow$.[94] Moreover, from $O! = [\lambda x\, \Diamond E!x]$ and $[\lambda x\, \Diamond E!x]\downarrow$ it will follow that $O!\downarrow$, by the symmetry of identity and the substitution of identicals; see (115.1).

By Convention (17.2.b), $O! =_{df} [\lambda y\, \Diamond E!y]$ is also an instance of the definition. Indeed, given this Convention, (.1) functions as the definition schema

[93] Some authors take $\tau\downarrow$ to represent '*definedness*'. But we're already using the technical term *defined* and loaded it with special meaning, both theoretically and metatheoretically, and so we won't use *definedness* in what follows.

[94] Since the matrix of $[\lambda x\, \Diamond E!x]$ contains no encoding formulas, no variable bound by the λ is in encoding position and so the definiens is a core λ-expression; see the definition of *core* in (9.2). So axiom (39.2) will guarantee that $[\lambda x\, \Diamond E!x]\downarrow$. Then our theory of definitions-by-identity (73) will ensure that since the definiens of (.1) is significant, the identity $O! = [\lambda x\, \Diamond E!x]$ is derivable.

$O! =_{df} [\lambda\nu\, \Diamond E!\nu]$. An axiom asserted in the next chapter, namely α-Conversion (48.1), will guarantee that $[\lambda x\, \Diamond E!x] = [\lambda y\, \Diamond E!y]$, given that $[\lambda x\, \Diamond E!x]\downarrow$ and the fact that $[\lambda x\, \Diamond E!x]$ and $[\lambda y\, \Diamond E!y]$ are alphabetic variants. So $[\lambda x\, \Diamond E!x]$ and $[\lambda y\, \Diamond E!y]$ will be substitutable for one another in any context, by the axiom for the substitution of identicals (41). This holds for any alphabetic variant of $[\lambda x\, \Diamond E!x]$. So it really doesn't matter which alphabetic variant of $[\lambda x\, \Diamond E!x]$ we use to define $O!$; it will be derivable that $O! = [\lambda\nu\, \Diamond E!\nu]$, for any individual variable ν.

We may also define *being abstract*, written $A!$, as follows:

(.2) $A! =_{df} [\lambda x\, \neg\Diamond E!x]$

I.e., *being abstract* is defined as *not possibly exemplifying concreteness*. Since it will be axiomatic that $[\lambda x\, \neg\Diamond E!x]\downarrow$ (39.2), this definition will yield the identity $A! = [\lambda x\, \neg\Diamond E!x]$ as a theorem. Remarks analogous to those following definition of $O!$ apply to $A!$.

(23) **Definitions:** Identity. As noted earlier, the definition of identity requires four cases. We first say that individuals are identical just in case either they are both ordinary and necessarily exemplify the same properties or they are both abstract and necessarily encode the same properties:

(.1) $x = y \equiv_{df} (O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)) \vee (A!x \,\&\, A!y \,\&\, \Box\forall F(xF \equiv yF))$

By our Convention (17.2), we obtain instances of this definition by uniformly replacing x and y by any individual terms κ_1 and κ_2, respectively, and by uniformly replacing F by any property variable Ω, provided Ω doesn't occur free in the terms replacing x and y. See Remark (35), where we examine the behavior of the above definition for the definiendum $x = \imath zGz$.

In the discussion accompanying the Encoding Formula Convention (17.3), we saw that x and y occur in encoding position in $x = y$. So the λ in $[\lambda xy\, x = y]$ binds variables in encoding position. The same applies to the λ in each of $[\lambda x\, x = y]$, $[\lambda x\, x = a]$, and $[\lambda y\, x = y]$, and $[\lambda y\, a = y]$. These are not core λ-expressions and therefore will *not* be among the terms guaranteed to have a denotation by axiom (39.2). This forestalls the McMichael/Boolos paradox discussed in Chapter 2, Section 2.1.

The second case for the definition of identity governs properties. For the following case, we suppress the superscript on F^1 and G^1 and so use F and G as unary relation variables under Convention (17.2). We then say that F *is identical to* G just in case F and G both exist and, necessarily, all and only the individuals that encode F encode G. Formally:

(.2) $F = G \equiv_{df} F\downarrow \,\&\, G\downarrow \,\&\, \Box\forall x(xF \equiv xG)$

See Remark (36) for an extended discussion of why the existence clauses have been added to the definiens. Though the variables F and G in the above function as metavariables under Convention (17.2), they function normally when we state axioms and theorems. So since it will be axiomatic that $F{\downarrow}$ and $G{\downarrow}$, by axiom (39.2), we will be able to derive, from the relevant instance of the above definition, that $F = G \equiv \Box\forall x(xF \equiv xG)$. We'll then be able to infer that this holds for all F and G. Given the negative free logic of complex terms that we'll adopt, the following universal closure of this theorem, namely $\forall F\forall G(F = G \equiv \Box\forall x(xF \equiv xG))$, can only be instantiated to significant property terms; see theorem (116.1). Contrast this with the definition displayed above, which can be instantiated to any property terms, significant or otherwise.

The third case of the definition of identity governs n-ary relations ($n \geq 2$); it defines relation identity in terms of the notions of existence and property identity. Let F and G be two n-ary relation variables ($n \geq 2$). Then we say that *F is identical to G* if and only if F and G both exist and each way of applying F and G to $n-1$ objects results in identical properties:

$$
\begin{aligned}
(.3)\ \ F = G \ \equiv_{df}\ & F{\downarrow} \ \&\ G{\downarrow}\ \& \\
& \forall y_1 \ldots \forall y_{n-1}([\lambda x\, Fxy_1 \ldots y_{n-1}] = [\lambda x\, Gxy_1 \ldots y_{n-1}]\ \& \\
& \quad [\lambda x\, Fy_1xy_2 \ldots y_{n-1}] = [\lambda x\, Gy_1xy_2 \ldots y_{n-1}]\ \& \ldots \& \\
& \quad [\lambda x\, Fy_1 \ldots y_{n-1}x] = [\lambda x\, Gy_1 \ldots y_{n-1}x])
\end{aligned}
$$

See Remark (37) for a discussion of why relation identity is not defined by using n-ary encoding to generalize the definition of property identity.

The fourth case of the definition of identity governs propositions and is stated in terms of existence and property identity. Where p and q are any 0-ary relation variables, we may say that p is identical to q just in case p and q both exist and *being such that* p is identical to *being such that* q:

$$(.4)\ \ p = q \ \equiv_{df}\ p{\downarrow} \ \&\ q{\downarrow} \ \&\ [\lambda x\, p] = [\lambda x\, q]$$

It might be of interest to note that since it will be axiomatic that $p{\downarrow}$ and $q{\downarrow}$ by axiom (39.2), we will be able to derive, from the relevant instance of the above definition, that $p = q \equiv [\lambda y\, p] = [\lambda y\, q]$. See theorem (116.2).

(24) **Definitions:** Non-identity. We now introduce the notion of *non-identity* in terms of *negation* and *identity* as follows:

$$\tau \neq \sigma \ \equiv_{df}\ \neg(\tau = \sigma),$$
provided τ and σ are both terms of the same type

This is not just a rewrite convention, for here we are defining *non-identity* in terms of primitive and defined notions. So we're not just introducing easier-to-read notation.

It is important to observe that we intentionally *omitted* the existence clauses $\tau\downarrow$ and $\sigma\downarrow$ from the definiens. That's because we want formulas of the form $\tau \neq \sigma$ to be provable not just when both τ and σ denote entities that fail to be identical but also when either τ or σ fails to denote. For example, if $\neg([\lambda x\, \psi]\downarrow)$, then by theorem (107.1), which asserts $\tau = \sigma \rightarrow \tau\downarrow$, it will follow that $\neg([\lambda x\, \psi] = [\lambda x\, \psi])$. So given the above definition, we may conclude $[\lambda x\, \psi] \neq [\lambda x\, \psi]$. Clearly, we are *not* preserving the suggestion that $\tau = \tau$ is true even when τ denotes nothing! Though some philosophers and logicians have argued that $\tau = \tau$ should always be true, the case is not compelling and so the suggestion will not be preserved in the present system.

(**25**) **Remark:** Four Final Observations About Definitions. The following observations about definitions may prove to be of interest. First, definitions often have a dual role. Sometimes they merely introduce a new technical expression that either (a) captures a certain condition about objects or relations (or both) or (b) abbreviates a complex term. But sometimes they provide an analysis of (one sense of) a logically or philosophically significant word or phrase of natural language. As examples of this latter role, (18.1) analyzes one sense of 'and', (18.4) analyzes one sense of 'there exists', and (18.5) analyzes one sense of 'possibly'. Similarly, definition (23.1) of '=' offers an analysis of the English expression 'is identical to' as it is used in claims such as "Hesperus is identical to Phosphorus" or "Clark Kent is identical to Superman". Definition (23.2), which provides a separate case of the definition of '=', offers an analysis of the phrase 'is identical to' as it is used in such claims as:

- *being a brother* is identical to *being a male sibling*
- *being a circle* is identical to *being a closed plane figure every point of which lies equidistant from some given point*

Of course, the hope is that such definitions provide insightful analyses that demonstrate the power of the primitive notions of the language. For example, the definition of property identity in (23.2) is introduced as part of our *theory* of identity: it (a) analyzes the identity of properties in terms of the primitives of our language, (b) provides a precise reconstruction of a notion thought to be mysterious, (c) is consistent with the intuition that properties may be distinct even if necessarily exemplified by the same objects, and (d) leads to a proof that the identity conditions of properties are extensional, i.e., leads to a proof that properties F and G are identical just in case they are encoded by the same objects. This is theorem (189).

Second, note that there are two ways to define new 0-ary terms. Consider:

$$q_0 \equiv_{df} \exists x(E!x \,\&\, \neg\mathcal{A}E!x)$$

$$q_0 =_{df} \exists x(E!x \,\&\, \neg\mathcal{A}E!x)$$

The second definition is well-formed because both terms flanking the symbol $=_{df}$ are 0-ary relation terms. Moreover, the definiens is provably significant – a theorem proved in Chapter 9, namely, (104.2), guarantees that the proposition $\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$ exists, i.e., that $(\exists x(E!x \,\&\, \neg\mathcal{A}E!x))\downarrow$. So the second definition will imply identities of the form $q_0 = \exists x(E!x \,\&\, \neg\mathcal{A}E!x)$.

Though only one definition of q_0 is allowed within the system, one may introduce it via either the definition-by-≡ or the definition-by-=. The choice ultimately rests on the inferential role that one wants q_0 to have. We'll formalize the distinct inferential roles of definitions-by-≡ and definitions-by-= in Chapter 9, in items (72) and (73). But, to a first approximation, we can say the choice of definition comes down to whether one wants to be able to derive a necessary equivalence or an identity between q_0 and $\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$. In a hyperintensional system such as ours, a necessary equivalence doesn't imply identity. We therefore leave the choice to be determined by the occasion and the way in which the definition is to be applied. In cases like q_0, however, we'll typically use the stronger definition-by-=, primarily so that we can simplify proofs.

Third, one of the axioms for quantification theory that we stipulate in the next chapter asserts, among other things, that the primitive constants of the language are significant; see (39.2). This axiom does *not* apply to constants that have been introduced by definition and thereby avoids the following problem. Our theory of definitions-by-identity will allow us to introduce a new constant, say d_{31}, by the definition $d_{31} =_{df} \imath x(Px \,\&\, \neg Px)$. Since $\imath x(Px \,\&\, \neg Px)$ will provably fail to have a denotation, it is not clear what use there is for such a definition. But, in this work, we aren't requiring that one *prove in advance* that the definiens is significant in order to introduce a definition-by-identity. As we shall see in Chapter 9, in item (73), the inferential role of a definition-by-identity is to make it axiomatically true that if the definiens is significant, then the identity holds, otherwise the definiendum fails to be significant. So, in the case we're describing, the definition introduces the following axiom:

$$(\imath x(Px \,\&\, \neg Px)\downarrow \to (d_{31} = \imath x(Px \,\&\, Px))) \,\&\, (\neg \imath x(Px \,\&\, \neg Px)\downarrow \to \neg d_{31}\downarrow)$$

So we can't use axiom (39.2) to assert $\tau\downarrow$ for *every* constant τ, since the definiens of a constant introduced by definition might fail to denote. A similar consideration applies to new relation constants defined in terms of paradoxical λ-expressions that aren't significant, such as $[\lambda x\, \exists G(xG \,\&\, \neg Gx)]$ (this expression is involved in the Clark/Boolos paradox). It will be provable that $\neg[\lambda x \exists G(xG \,\&\, \neg Gx)]\downarrow$ (192.1). While our theory of definition-by-identity allows one to introduce $S_{32} =_{df} [\lambda x\, \exists G(xG \,\&\, \neg Gx)]$, axiom (39.2) does not thereby sanction $S_{32}\downarrow$, since S_{32} is a constant introduced by a definition.[95]

[95]Note how our discussion applies to (a) the case where one defines d_{31} as above and then defines $P_{33} =_{df} [\lambda x\, Rxd_{31}]$, and (b) the case where one defines the term S_{32} as above and then defines

The fourth and final observation is that the discussion of definitions in this section completes the precise syntactic characterization of our formalism. This forestalls the concerns that Gödel (1944, 120) raised about Whitehead and Russell's *Principia Mathematica*:

> It is to be regretted that this first comprehensive and thorough-going presentation of a mathematical logic and the derivation of mathematics from it [is] so greatly lacking in formal precision in the foundations (contained in *1–*21 of *Principia*) that it presents in this respect a considerable step backwards as compared with Frege. What is missing, above all, is a precise statement of the syntax of the formalism. Syntactical considerations are omitted even in cases where they are necessary for the cogency of the proofs, in particular in connection with the "incomplete symbols".

It is important to have taken some pains in the above development to ensure that such a criticism doesn't apply to the present effort. Attention to the syntactic details should also be useful to those interested in both the computational implementation and metatheoretic properties of the system.

7.3 Explanatory Remarks: Digression

In this section, we develop a number of remarks that will offer some perspective on, and examples of, the metadefinitions and definitions in this chapter. Unfortunately, we shall need to appeal occasionally to facts about our system not yet in evidence. Some of the choices in the foregoing can't be explained adequately without mentioning developments (definitions, axioms, and theorems) that are formulated much later in the text. Nevertheless, I've tried to make the discussion intuitive and, in some cases, have repeated later developments so that most of the Remarks can be read on their own.

(26) **Remark:** About Alphabetic Variants. *Alphabetically-variant* formulas and terms are complex expressions that differ only orthographically with respect to the *bound* variables they contain and intuitively have the same meaning. Thus, Fx and Fy are *not* alphabetic variants, nor are Fa and Ga. Rather, in the simplest cases:

- $\forall xFx$ and $\forall yFy$ are alphabetically-variant formulas; they not only have the same truth conditions but denote the same proposition, namely, that every individual exemplifies F.

$d_{33} = \imath x\varphi$, for some formula φ in which S_{32} occurs. In these cases, our theory of definitions-by-identity will guarantee that P_{33} and d_{33} are significant if their definientia are.

- $\imath xFx$ and $\imath yFy$ are alphabetically-variant descriptions; they either both denote the unique individual that in fact exemplifies F if there is one, or both denote nothing if there isn't.
- $[\lambda x\, \neg Fx]$ and $[\lambda y\, \neg Fy]$ are alphabetically-variant relation terms; they denote the same relation.

However, we shall need to define the notion of *alphabetic variant* for formulas and terms of arbitrary complexity. So we need a definition on which the following pairs of expressions count as alphabetic variants:

- $\forall F(Fx \equiv Fy)$ / $\forall G(Gx \equiv Gy)$
- $\imath x\forall yMyx$ / $\imath y\forall zMzy$
- $[\lambda y\, Ry\imath zQz]$ / $[\lambda z\, Rz\imath xQx]$
- $[\lambda\, Pa \to \forall F\, Fa]$ / $[\lambda\, Pa \to \forall G\, Ga]$
- $[\lambda\, [\lambda y\, \neg Fy]a \to \forall xMx]$ / $[\lambda\, [\lambda z\, \neg Fz]a \to \forall yMy]$

Once we have a definition that counts the above pairs as alphabetic variants, we'll be in a position to state and prove a number of theorems, such as (1) alphabetically variant formulas denote the same proposition (111.4), (2) alphabetically-variant formulas are equivalent (111.5), (3) alphabetically-variant definite descriptions have the same denotation if they have one (154), (4) a formula is derivable from some premises if and only if its alphabetic variants are derivable from those premises (114), and (5) a formula is a theorem if and only if its alphabetic variants are theorems, which is a special case of (114).

The definition of alphabetic variance also enables us to stipulate an axiom schema that intuitively guarantees that alphabetically-variant λ-expressions denote the same relation if they have a denotation. This is the schema α-Conversion, introduced in (48.1). It has the following as instances, where $\downarrow$ is defined in (20.2) and (20.3) and intuitively asserts that the expression to which it attaches has a denotation:

$$[\lambda y\, Ry\imath zQz]\!\downarrow \to [\lambda y\, Ry\imath zQz] = [\lambda x\, Rx\imath yQy]$$

$$[\lambda\, Pa \to \forall F\, Fa]\!\downarrow \to [\lambda\, Pa \to \forall F\, Fa] = [\lambda\, Pa \to \forall G\, Ga]$$

$$[\lambda\, [\lambda y\, \neg Fy]a \to \forall xMx]\!\downarrow \to$$
$$[\lambda\, [\lambda y\, \neg Fy]a \to \forall xMx] = [\lambda\, [\lambda z\, \neg Fz]a \to \forall yMy]$$

Note that the second and third conditionals introduce an equation between two 0-ary relation terms.

Our definition of alphabetic variance begins with clause (16.1), which introduces *linked* and *independent* variables. Here are some examples:

- In the formula $\forall F(Fa \equiv Fb)$, each occurrence of the variable F is linked to every other occurrence.
- In the formula $\forall F Fa \equiv \forall F Fb$, the first two occurrences of F are linked to each other and the last two occurrences of F are linked to each other, while each of the first two occurrences is independent of each of the last two occurrences and vice versa.
- In the formula $\forall x Fx \rightarrow Fx$, the first two occurrences of x are linked, and both are independent of the third, given that this formula is shorthand for $(\forall x Fx) \rightarrow Fx$.
- In the term $\imath x(Fx \rightarrow Gy)$, the two occurrences of x are linked.
- In the term $[\lambda y\, \forall x Gx \rightarrow (Gx \,\&\, Gy)]$, the two occurrences of y are linked (both are bound by the λ), the first two occurrences of x are linked, and both of those occurrences of x are independent of the third occurrence of x. Also, all three occurrences of G are free and hence linked.

Since *linked* is an equivalence condition on variables in a term or formula, we call the cells of each partition a linkage group. In the first example above, there is one linkage group for F, and in the fourth example, there is one linkage group for x. In the second example, there are two linkage groups for F and in the third example, there are two linkage groups for x. In the fifth example, there are two linkage groups for x, one linkage group for y, and one linkage group for G.

In (16.2), we defined BV-notation. Some examples help illustrate this definition:

- When $\varphi = \forall F(Fx \equiv Fy)$, then φ in BV-notation is $\varphi[F,F,F]$
- When $\tau = \imath x \forall y Myx$, then τ in BV-notation is $\tau[x,y,y,x]$
- When $\tau = [\lambda y\, Ry\imath xQx]$, then τ in BV-notation is $\tau[y,y,x,x]$
- When $\tau = [\lambda\, Pa \rightarrow \forall F Fa]$, then τ in BV-notation is $\tau[F,F]$
- When $\tau = [\lambda x\, \neg Fx]a \rightarrow \forall y My$, then τ in BV-notation is $\tau[x,x,y,y]$

In definitions (16.3)– (16.5), we completed the definition of *alphabetic variant*. The following examples of alphabetically-variant formulas and terms may also prove useful:

- $\varphi = \forall F Fa \equiv \forall G Gb$
  $\varphi' = \forall F Fa \equiv \forall F Fb$
- $\varphi = \forall x Rxx \rightarrow \exists y Syz$
  $\varphi' = \forall y Ryy \rightarrow \exists x Sxz$

- $\tau = \imath z(Fz \rightarrow Gy)$
  $\tau' = \imath x(Fx \rightarrow Gy)$

- $\tau = [\lambda y\, \forall x Gx \rightarrow (Gz \,\&\, Gy)]$
  $\tau' = [\lambda x\, \forall y Gy \rightarrow (Gz \,\&\, Gx)]$

There are a number of final observations to make about the definition of *alphabetic variant*. First, note that even though all the free occurrences of a variable α in φ (or τ) are linked, they are preserved as is (i.e., without change and in the same position) in any alphabetic variant φ' (or τ') since the free occurrences are not listed in BV-notation.

Second, note that our definitions imply that:[96]

> If τ is a term occurring in φ and τ' is an alphabetic variant of τ, then if φ' is the result of replacing one or more occurrences of τ in φ by τ', then φ' is an alphabetic variant of φ.
>
> If φ is a formula occurring in τ and φ' is an alphabetic variant of φ, then if τ' is the result of replacing one or more occurrences of φ in τ by φ', then τ' is an alphabetic variant of τ.

Here are some example pairs of the preceding facts (the first two pairs are examples of the first fact and the second two pairs are examples of the second fact):

- Where $\varphi_1 = \forall F(F\imath xPx \equiv Fb)$, $\tau_1 = \imath xPx$, and $\tau_1{}' = \imath yPy$, then $\forall F(F\imath yPy \equiv Fb)$ is an alphabetic variant of φ_1

- Where $\varphi_2 = \forall x([\lambda y\, \neg Fy]x \equiv \neg Fx)$, $\tau_2 = [\lambda y\, \neg Fy]$, and $\tau_2{}' = [\lambda z\, \neg Fz]$, then $\forall x([\lambda z\, \neg Fz]x \equiv \neg Fx)$ is an alphabetic variant of φ_2

- Where $\tau_3 = [\lambda y\, \forall x Gx \rightarrow Gy]$, $\varphi_3 = \forall x Gx$, and $\varphi_3{}' = \forall z Gz$, then $[\lambda y\, \forall z Gz \rightarrow Gy]$ is an alphabetic variant of τ_3

- Where $\tau_4 = \imath y \forall F(Fy \equiv Fa)$, $\varphi_4 = \forall F(Fy \equiv Fa)$, and $\varphi_4{}' = \forall G(Gy \equiv Ga)$, then $\imath y \forall G(Gy \equiv Ga)$ is an alphabetic variant of τ_4

In the first case, $\imath yPy$ is an alphabetic variant of $\imath xPx$, and replacing the latter by the former in φ_1 yields $\varphi_1{}'$. In this case, φ_1 in BV-notation is $\varphi[F,F,x,x,F]$ and $\varphi_1{}'$ in BV-notation is $\varphi_1{}'[F,F,y,y,F]$. So (a) $\varphi_1{}' = \varphi_1[F/F,F/F,y/x,y/x,F/F]$,

[96] The first of the following two facts holds because all the bound variable occurrences in $\tau[\alpha_1,\ldots,\alpha_n]$ will appear in $\varphi[\gamma_1,\ldots,\gamma_m]$ and the replacements for the γ_is needed to replace τ with $\tau'[\beta_1,\ldots,\beta_n]$ can not break any linkage groups. The only variable occurrences in τ that could become linked to non-τ variable occurrences in φ are the free variables of τ and those cannot be changed in τ' as noted above. The same reasoning explains why the second of the following two facts holds.

and (b) any two variables in the list of bound variables for φ_1 are linked in φ_1 if and only the corresponding variables in the list of bound variables for ${\varphi_1}'$ are linked in ${\varphi_1}'$. We leave the explanation of the remaining cases as exercises for the reader.

Third, note that our definitions give rise to the following fact:

Metatheorem ⟨7.6⟩
Alphabetic variance is an equivalence condition on the complex formulas of our language, i.e., alphabetic variance is reflexive, symmetric, and transitive.

A proof can be found in the Appendix to this chapter.

A final observation concerns an important metatheorem that can be approached by way of an example. Let φ be a formula of the form $\neg\psi$. Then any alphabetic variant of φ we pick will be a formula of the form $\neg(\psi')$, for some alphabetic variant ψ' of ψ. For example, if φ is $\neg\forall xFx$, so that ψ is $\forall xFx$. Then pick any alphabetic variant of φ, say, $\neg\forall yFy$. In this case, the formula $\forall yFy$ is the witness ψ' such that φ' is $\neg(\psi')$. This generalizes to formulas of arbitrary complexity, so that we have:

Metatheorem ⟨7.7⟩: Alphabetic Variants of Complex Formulas and Terms.

(a) If φ is a formula of the form $[\lambda\,\psi]$ ($\neg\psi$, $\mathcal{A}\psi$, or $\Box\psi$), then each alphabetic variant φ' is a formula of the form $[\lambda\,\psi']$ ($\neg(\psi')$, $\mathcal{A}(\psi')$ or $\Box(\psi')$, respectively), for some alphabetic variant ψ' of ψ.

(b) If φ is a formula of the form $\psi \to \chi$, then each alphabetic variant φ' is a formula of the form $\varphi' \to \psi'$, for some alphabetic variants φ' and ψ', respectively, of φ and ψ.

(c) If φ is a formula of the form $\forall\alpha\psi$, then each alphabetic variant φ' is a formula of the form $\forall\beta(\psi'^{\beta}_{\alpha})$, for some alphabetic variant ψ' of ψ and some variable β substitutable for α in ψ' and not free in ψ'.[97]

(d) If τ is a term of the form $\imath\nu\varphi$, then each alphabetic variant τ' is a term of the form $\imath\mu(\varphi'^{\mu}_{\nu})$, for some alphabetic variant φ' of φ and individual variable μ substitutable for ν in φ' and not free in φ'.

(e) If τ is a term of the form $[\lambda\nu_1 \dots \nu_n\,\varphi]$ ($n \geq 1$), then each alphabetic variant τ' is a term of the form $[\lambda\mu_1 \dots \mu_n\,(\varphi'^{\mu_1,\dots,\mu_n}_{\nu_1,\dots,\nu_n})]$, for some alphabetic variant φ' of φ and individual variables μ_i ($1 \leq i \leq n$) substitutable, respectively, for the ν_i in φ' and not free in φ').[98]

In the Appendix, the proof of this metatheorem is left as an exercise.

[97] Note that this allows for the case where β is α or the case where ψ' is ψ.

[98] Note that this allows for the case where $\mu_1,\dots,\mu_n$ are, respectively, $\nu_1,\dots,\nu_n$.

(27) **Remark:** Why Free Variables in Definitions Should Be, or Function As, Metavariables. In our system, we cannot use object-language variables to formulate definitions if those object-language variables are given their normal interpretation and the definition is understood classically. To see the problem with respect to free variables, let's first consider a definition-by-$\equiv$. Suppose one wanted to define the condition, object x *contingently exemplifies* property F, by stipulating that it holds just in case x exemplifies F but doesn't necessarily exemplify F. One might expect to see this definition formalized using object language variables as follows:

(A) $\mathit{ContingentlyExemplifies}(x,F) \equiv_{df} Fx \;\&\; \neg\Box Fx$

In the present context, however, definitions like the above are problematic and have to be replaced by a definition schema. The traditional understanding of definitions like (A) fails for systems such as ours, in which individual and relation terms may fail to denote.

On the traditional understanding, a definition such as (A) becomes available to the deductive system as a biconditional axiom, i.e., as an axiom where $\equiv$ replaces $\equiv_{df}$ in (A). That is its traditional inferential role. This understanding of definitions suffices in simple systems of classical logic, in which every term of the language has a denotation and even the free variables that may occur in an asserted formula have some (arbitrarily assigned) value. The classical logic of quantification permits the instantiation of *any* individual or relation term τ of such a language into a universally quantified claim of the form $\forall\alpha\varphi$.

So in such systems, (A) not only extends the language with new formulas of the form $\mathit{ContingentlyExemplifies}(\kappa,\Pi)$ (where κ is any individual term and Π any unary relation term), but also extends the deductive system with new axioms, including both of the following:[99]

(B) $\mathit{ContingentlyExemplifies}(x,F) \equiv (Fx \;\&\; \neg\Box Fx)$

(C) $\forall x\forall F(\mathit{ContingentlyExemplifies}(x,F) \equiv (Fx \;\&\; \neg\Box Fx))$

In classical systems, every object term κ and every property term Π can be instantiated, respectively, for $\forall x$ and $\forall F$ in (C) to yield biconditional theorems stating the necessary and sufficient conditions for $\mathit{ContingentlyExemplifies}(\kappa,\Pi)$.

[99]In some systems (not ours, though), where both (a) formulas with free variables are assertible and (b) the Rule of Generalization (GEN) is primitive, (A) implicitly adds only (B) as an axiom, since (C) follows from (B) by GEN. In other systems (again, not ours), where formulas with free variables are not assertible, the definition implicitly adds only (C) as an axiom. As we shall see, though, in our system, where (a) formulas with free variables are assertible but (b) the Rule of Generalization (GEN) is a *metarule*, the definition implicitly adds both (B) and (C) as axioms, since all of the *closures* of (B), including universal closures, are taken as axioms. We'll make this explicit in Chapters 8 and 9.

However, our system includes both complex individual terms (definite descriptions) and complex n-ary ($n \geq 1$) relation terms (λ-expressions) that may fail to have a denotation.[100] The *negative free logic* that governs such complex terms ensures that a term τ can't be instantiated into a universal claim unless it is (provably) significant. Now consider a description like $\imath z(Pz \,\&\, \neg Pz)$, which will provably fail to have a denotation.[101] Let's abbreviate $\imath z(Pz \,\&\, \neg Pz)$ as $\imath z\psi$. If definition (A) implicitly introduces the axiom (C), then although the classical logic of constants would allow us to instantiate the relation constant P for the universal quantifier $\forall F$ in (C), the *negative free logic* of non-denoting terms would not allow us to instantiate the description $\imath z\psi$ for the universal quantifier $\forall x$ in (C). Thus, we wouldn't be able to derive from (C):

(D) $\mathit{ContingentlyExemplifies}(\imath z\psi, P) \equiv (P\imath z\psi \,\&\, \neg\Box P\imath z\psi)$

So given the classical understanding of definitions on which (A) implicitly introduces (B) and (C) as axioms, negative free logic prevents us from deriving (D). Thus, the notion *ContingentlyExemplifies* isn't really defined when $\imath z\psi$ fails to denote; no biconditional theorem states the necessary and sufficient conditions under which $\mathit{ContingentlyExemplifies}(\imath z\psi, \Pi)$, for any term Π.

Similarly, let Π be a property term that provably fails to denote, such as one of the λ-expressions that leads to the Clark/Boolos paradox. For simplicity, let $[\lambda x\, \varphi]$ be an arbitrary such property term. Then, analogously, if (A) implicitly introduces the axiom (C), then although the classical logic of constants would allow us to instantiate the individual constant a for the quantifier $\forall x$ in (C), the *negative free logic* of non-denoting terms would not allow us to instantiate the λ-expression $[\lambda x\, \varphi]$ for the quantifer $\forall F$ in (C). Thus, we wouldn't be able to obtain the following as a theorem:

(E) $\mathit{ContingentlyExemplifies}(a, [\lambda x\, \varphi]) \equiv ([\lambda x\, \varphi]a \,\&\, \neg\Box[\lambda x\, \varphi]a)$

And, clearly, if both $\imath z\psi$ and $[\lambda x\, \varphi]$ fail to have a denotation, we wouldn't be able to obtain the following as a theorem:

(F) $\mathit{ContingentlyExemplifies}(\imath z\psi, [\lambda x\, \varphi]) \equiv ([\lambda x\, \varphi]\imath z\psi \,\&\, \neg\Box[\lambda x\, \varphi]\imath z\psi)$

[100] We don't include 0-ary relation terms (and thus formulas) in this list, since it will be a theorem that all such terms have a denotation; see theorem (104.1).

[101] In our system, the notion of a term's having a denotation is represented in the object language by the logico-metaphysical notion of *existence*, which is defined in (20) and symbolized by $\downarrow$. So the proof that $\imath z(Pz \,\&\, \neg Pz)$ fails to have a denotation, i.e., the proof that $\neg\imath z(Pz \,\&\, \neg Pz)\downarrow$, goes by way of reductio. For suppose $\imath z(Pz \,\&\, \neg Pz)\downarrow$. Then by definition (20.1), there is some property, say Q, such that $Q\imath z(Pz \,\&\, \neg Pz)$. This implies, by the laws of definite descriptions, that $\exists z(Pz \,\&\, \neg Pz)$, which yields a contradiction for any witness to the existence claim. But in our system, $Q\imath z(Pz \,\&\, \neg Pz)$ immediately implies $\imath z(Pz \,\&\, \neg Pz)\downarrow$, by axiom (39.5.a), thus directly contradicting our reductio assumption.

So the fact that (D), (E), and (F) aren't theorems means that definition (A) isn't sufficient to guarantee that following formulas have the truth conditions specified by the definition:

$$\textit{ContingentlyExemplifies}(\imath z\psi, P)$$

$$\textit{ContingentlyExemplifies}(a, [\lambda x\, \varphi])$$

$$\textit{ContingentlyExemplifies}(\imath z\psi, [\lambda x\, \varphi])$$

We can avoid this problem by using metavariables and formulating a definition schema. Let κ be a metavariable ranging over individual terms and Π be a metavariable ranging over unary relation terms. Then (G) avoids the problems just noted (A):

(G) $\textit{ContingentlyExemplifies}(\kappa, \Pi) \equiv_{df} \Pi\kappa \,\&\, \neg\Box\Pi\kappa$

A definition schema such as (G) implicitly extends our language with the new syncategorematic expression *ContingentlyExemplifies* and new formulas of the form $\textit{ContingentlyExemplifies}(\kappa, \Pi)$. But, just as importantly, (G) will implicitly introduce (the closures of) the instances of the following schema as new axioms:

$$\textit{ContingentlyExemplifies}(\kappa, \Pi) \equiv (\Pi\kappa \,\&\, \neg\Box\Pi\kappa)$$

Given such an understanding of definition schemata, (G) yields (D), (E), and (F) as theorems, since these are all instances of the above biconditional. Consequently, the use of metavariables in (G) is required and (A), strictly speaking, doesn't suffice as a definition. However, since (G) is more complex and more difficult to read than (A), we have formulated Convention (17.2.a), which specifically allows the free variables in (A) to function as metavariables, so that (D), (E), and (F) become instances of the definition.

Now let's consider definitions-by-=. Though object-language variables that may occur free in a definition-by-= should also function as metavariables, it is *not* for the reason just outlined. To see why, let's consider a somewhat concocted example that has some interesting probative features. Consider theorem (177.1), i.e., $\imath x(x = y)\downarrow$, which asserts, for an arbitrary object y: the x such that x is identical to y exists. Let's use $\imath x(x = y)$ as the definiens for ι_y in the following definition-by-=:

(ϑ) $\iota_y =_{df} \imath x(x = y)$

For example, ι_a ('*the a*') is thus defined as the individual x identical to a. (ϑ) is a fine definition given that the definiens has a denotation for each value assigned to the free variable y. No matter what is assigned to y, ι_y denotes the individual that is identical to y, i.e., denotes y.

Traditionally, (ϑ) would be understood as extending our language with a host of new complex terms. Though (ϑ) uses the free object-language variable y, it is standard to assume that (ϑ) would extend our language with terms of the form ι_κ, where κ is any term. So all of the following would be well-formed: ι_y, $\iota_{\imath z\varphi}$, $\iota_{\imath x\psi}$, ι_{ι_y}, etc.

Also traditionally, (ϑ) would be understood as implicitly introducing the closures of the axiom:

(ξ) $\iota_y = \imath x(x=y)$

In a classical logic, in which all terms have denotations, (ξ) yields an identity axiom for every term other than x, only terms other than x are substitutable for y in the matrix of the universal closure $\forall y(\iota_y = \imath x(x=y))$

At first glance, this understanding of the inferential role of definitions-by-= would appear to be desirable, for in our system, we have individual terms that fail to denote; our negative free logic does not permit us to instantiate empty terms into $\forall y(\iota_y = \imath x(x=y))$, for that would yield identities in which the terms flanking the identity sign fail to have a denotation. So we don't want to interpret (ϑ) as introducing, for any individual term κ, the axiom schema:

$$\iota_\kappa = \imath x(x=\kappa)$$

If $\imath z\psi$ is a description that provably fails to have a denotation, the above would yield the following as an axiom:

(ζ) $\iota_{\imath z\psi} = \imath x(x=\imath z\psi)$

In (ζ), both terms flanking the identity sign fail to have a denotation, and such a claim can't be axiomatic.[102] Identity statements can't be true when the terms flanking them are empty unless heroic measures are taken, something we'll forego here. So, in a negative free logic, the classical interpretation of (ϑ) as introducing (ξ) blocks the introduction of identities like (ζ) with non-denoting descriptions.

[102]Indeed, in our system, (ζ) would be provably false when $\imath z\psi$ fails to have a denotation, though to see this we have to cite definitions, axioms, and theorems not yet introduced. Assume $\neg(\imath z\psi\downarrow)$. Then choose a variable, say x, that isn't free in $\imath z\psi$. It follows by theorem (107.2) that $\neg(x=\imath z\psi)$. Since x isn't free in our assumption, it follows by GEN that this holds for any object x, i.e., that $\forall x\neg(x=\imath z\psi)$, i.e., $\neg\exists x(x=\imath z\psi)$.

From this last conclusion, we may infer, by the laws of definite descriptions and the definition of $\downarrow$ that $\neg\imath x(x=\imath z\psi)\downarrow$, as follows:

> Assume, for reductio, that $\imath x(x=\imath z\psi)\downarrow$. Then by definition of $\downarrow$ (20.1), it follows that for some property, say P, that $P\imath x(x=\imath z\psi)$. But, then, by Russell's analysis of descriptions (143)★, it follows *a fortiori* that that $\exists x(x=\imath z\psi)$, which contradicts what we proved above.

From $\neg\imath x(x=\imath z\psi)\downarrow$, we can again conclude, by theorem (107.2), that $\neg(\iota_{\imath z\psi} = \imath x(x=\imath z\psi))$. Thus, we have a proof of the negation of (ζ), when $\imath z\psi$ fails to be significant.

Thus, one might conclude at this point that we *should* interpret (ϑ) as introducing (ξ) and that we need not interpret the object-language variables in (ϑ) as metavariables. But this would be a mistake, for the classical interpretation doesn't yield a mechanism for proving that $\iota_{\imath z\psi}$ fails to have a denotation when $\imath z\psi$ fails to have a denotation! Though we saw in footnote 102 that $\neg\imath x(x = \imath z\psi)\downarrow$ is a consequence of $\neg\imath z\psi\downarrow$, ($\vartheta$) doesn't allow us to derive $\neg\iota_{\imath z\psi}\downarrow$ from $\neg\imath x(x = \imath z\psi)\downarrow$. Intuitively, if the definiens $\imath x(x = \imath z\psi)$, fails to be significant, then we should be able to derive that the definiendum $\iota_{\imath z\psi}$ fails to be significant. So the problem is that (ϑ), under the standard interpretation of its object-language variables, doesn't give us a means to conclude $\neg\iota_{\imath z\psi}\downarrow$ when $\neg\imath x(x = \imath z\psi)\downarrow$. The term $\iota_{\imath z\psi}$ is *well-formed* and we know the claim $\iota_{\imath z\psi}\downarrow$ to be false when $\imath x(x = \imath z\psi)$ is empty, but we have no means of proving it.

Our solution will be to let the object-language variables in definitions-by-= function as metavariables but *revise* our understanding of the inferential role of these definitions. We'll allow any terms to be substituted for the free object-language variables so that we have instances of the definition for every individual term of the language. But the inferential role of the definition, as formulated in (73), will be that it becomes a metarule stipulating that (the closures of) a certain conjunction of conditionals is (are) axiomatic. For the particular instance of (ϑ) we're now considering, namely, $\iota_{\imath z\psi} =_{df} \imath x(x = \imath z\psi)$, the metarule will stipulate that the following is a necessary axiom schema: if $\imath x(x = \imath z\psi)$ is significant, then $\iota_{\imath z\psi} = \imath x(x = \imath z\psi)$, and if $\imath x(x = \imath z\psi)$ fails to be significant, then $\iota_{\imath z\psi}$ fails to be significant.[103]

We'll discuss the inferential role of definitions-by-= in more detail in (73), (119), (120), and especially in Remark (283). In that discussion, we'll see that the object-language variables that occur free in such definitions still function as metavariables. So Convention (17.2.a) will apply even to definitions-by-=.

(**28**) **Remark:** Why Bound Variables in Definientia Should Be, or Function As, Metavariables. Consider definitions-by-≡ first. Suppose one wanted to define: x and y are *indiscernible* just in case x and y exemplify the same properties. We might introduce this definition formally as:

(ϑ) $Indiscernible(x, y) \equiv_{df} \forall F(Fx \equiv Fy)$

However, the following would be a perfectly good instance of the definition:

$$Indiscernible(x, y) \equiv_{df} \forall G(Gx \equiv Gy)$$

[103] It is important to note that our example has been over-simplified, since in this instance, the definiens $\imath x(x = \imath z\psi)$ is significant if and only if its argument $\imath z\psi$ is. But this is an artifact of the particular example. We'll see cases later where (i) the definiens doesn't have a denotation even though all of its argument terms do, and (ii) the definiens has a denotation even though one or more of its argument terms do not. See the examples discussed in Remark (283). For now, however, the example we've used suffices for the purpose of giving an overview explanation of the inferential role of definitions-by-=.

Though we shall eventually be able to prove that the alphabetically-variant formulas $\forall F(Fx \equiv Fy)$ and $\forall G(Gx \equiv Gy)$ are equivalent, this theorem doesn't become provable until (114). Moreover, we may not produce an instance of (ϑ) by substituting $\imath zFz$ for x, since $\imath zFz$ isn't substitutable for x in the definiens. So, if we want $Indiscernible(\imath zFz, y)$ to be defined, then something like the following has to be an instance of the definition:

$$Indiscernible(\imath zFz, y) \equiv_{df} \forall G(G\imath zFz \equiv Gy)$$

So, if we put aside the free variables (which we already know function as metavariables), it is important to interpret (ϑ) as follows, where Ω is a metavariable ranging over unary relation variables:

$$Indiscernible(x, y) \equiv_{df} \forall\Omega(\Omega x \equiv \Omega y)$$

But, clearly, this latter is more difficult to read and grasp, and so the Convention we formulate in (17.2.b) will allow us to suppose that the variables in the definiens of (ϑ) function as metavariables.

To see an example that shows why bound variables in the definientia of definitions-by-= should function as metavariables, we invoke an instance of a theorem schema proved much later (250), namely, that there is a *unique* abstract object that encodes all and only non-self-identical properties. Using the *uniqueness* quantifier $(\exists!x)$ defined in item (127.1), this theorem is formally respresented as:

$$\exists!x(A!x \,\&\, \forall F(xF \equiv F \neq F))$$

Consequently, it will be a theorem (252) that *the* abstract object that encodes exactly the non-self-identical properties exists, i.e.,

$$\imath x(A!x \,\&\, \forall F(xF \equiv F \neq F))\downarrow$$

So we may introduce the individual constant $\boldsymbol{a}_\varnothing$ as follows:[104]

$$\boldsymbol{a}_\varnothing \;=_{df}\; \imath x(A!x \,\&\, \forall F(xF \equiv F \neq F))$$

However, the following would also be a perfectly good instance of the definition:

$$\boldsymbol{a}_\varnothing \;=_{df}\; \imath y(A!y \,\&\, \forall G(yG \equiv G \neq G))$$

Though we will eventually prove that the alphabetic variants of a description that has a denotation all denote the same individual, this theorem doesn't

[104] Though we won't actually use this definition in our work, a definition such as this will be discussed in Remark (267). The new symbol $\boldsymbol{a}_\varnothing$ is a complex individual term with no free variables and should be considered as a single unit, though we sometimes use the boldface lowercase symbol $\boldsymbol{a}$ with a different decoration to introduce other new individual constants.

become provable until (154). So, until then, it important to ensure that the above definitions are interpreted as the following, where no special provisos are needed:[105]

$$\boldsymbol{a}_\varnothing =_{df} \imath\nu(A!\nu \,\&\, \forall\Omega(\nu\Omega \equiv \Omega \neq \Omega))$$

Clearly, the definitions of $\boldsymbol{a}_\varnothing$ with bound object-language variables in the definiens are easier to read and grasp than the one involving metavariables, and so when developing definitions of this kind, we'll use object-language variables under Convention (17.2).

The foregoing therefore explains (a) why the bound variables in the definientia of both definitions-by-≡ and definitions-by-= should be metavariables, and (b) why we have developed our Convention (17.2.b), which allows bound object-language variables to function as metavariables whenever that makes the definition easier to read and grasp.

(**29**) **Remark:** Why We Could Have Used Object-Language Variables in Definition (18). Though the reasons why it is actually clearer to use metavariables were given in (18), note that we could have invoked Convention (17.2.a) so as to formulate the definitions in (18) with object-language variables as follows:

(.1) $p \,\&\, q \equiv_{df} \neg(p \rightarrow \neg q)$

(.2) $p \vee q \equiv_{df} \neg p \rightarrow q$

(.3) $p \equiv q \equiv_{df} (p \rightarrow q) \,\&\, (q \rightarrow p)$

(.4) (.a) $\exists x p \equiv_{df} \neg\forall x \neg p$

(.b) $\exists F^n p \equiv_{df} \neg\forall F^n \neg p$ $\quad (n \geq 0)$

(.5) $\Diamond p \equiv_{df} \neg\Box\neg p$

By Convention (17.2.a), instances of these definitions are obtained by uniformly substituting 0-ary terms for p and q. Since our BNF (4) stipulates that formulas are 0-ary relation terms, we obtain instances of the above when any formulas are uniformly substituted for p and q. That's precisely the effect that the Greek metavariables φ and ψ have in definition schemata (18.1) – (18.5). Moreover, by Convention (17.2.b), instances of (.4.a) and (.4.b) above can be obtained by replacing the definiens with any alphabetic variant. That's precisely the effect of interpreting x in (.4.a) and F^n in (.4.b), respectively, as the metavariables ν and Π^n. So, combining all these conventions, (.4.a) and (.4.b) would be interpreted as:

[105] The instances of ν and Ω are variables and so ν can't occur free in Ω and Ω can't occur free in ν. So although, in the definiens, ν occurs under the scope of $\forall\Omega$ and Ω falls under the scope of $\imath\nu$, no provisos are needed; no variable can get captured when we uniformly replace ν by an individual variable and uniformly replace Ω by a unary relation variable.

$$\exists\nu\,\varphi \equiv_{df} \neg\forall\nu\neg\varphi$$

$$\exists\Pi^n\,\varphi \equiv_{df} \neg\forall\Pi^n\neg\varphi$$

But this reveals another reason why it is clearer to use metavariables as we did in (18): since α is a metavariable that ranges over both individual variables and n-ary relation variables ($n \geq 0$), these last two definitions can be captured by the single definition schema, namely, (18.4). This single definition schema has, as instances, all of the instances of the above two definition schemata.

(**30**) **Remark:** Concerning the Definition of Existence. When we assert, using the logical sense of 'existence', that 'Obama exists', 'The present king of France exists', or 'the property *being possibly concrete* exists', we semantically imply that the terms 'Obama', 'the present king of France', or 'being possibly concrete' denote or signify something. So the claim $\ulcorner\tau$ *exists*$\urcorner$, for any term τ, has ontological significance when true. Such claims may be false in languages containing either individual terms or relation terms that fail to denote. (For those who are reading this remark without having read footnote 92, the Quinean corner quotes serve to make it clear that the claim τ *exists* is not a metalinguistic assertion about τ, but an object-theoretic assertion that *uses* the term of our language which serves as the value of τ.)

In a classical, second-order language without identity but with non-denoting terms, one can't use the exemplification mode of predication to define logical existence generally for *both* individuals and relations. Exemplification serves well enough to define existence for individuals, but not well enough to define existence for relations. Here is why.

In a classical, second-order language, one may define: individual κ *exists* just in case κ exemplifies some property. Such a definition might be offered on the basis of the following claims, which both seem true:

> Obama *exists* if and only if Obama exemplifies some property.
>
> The present king of France *exists* if and only if the present king of France exemplifies some property.

These last two claims might be represented formally as follows:

$$o\!\downarrow\; \equiv \exists F Fo$$

$$\imath xKx\!\downarrow\; \equiv \exists F F\imath xKx$$

To see that these serve well for defining the existence of individuals, consider the second example. If for some property, say P, the exemplification formula $P\imath xKx$ is true, then the term '$\imath xKx$' denotes something. That's because a classical, second-order language is intuitively grounded on the principle that:

> An exemplification formula of the form $\Pi^n\kappa_1\ldots\kappa_n$ ($n \geq 1$) is true if and only if the terms Π^n, κ_1, ..., and κ_n all have a denotation and the individuals denoted by the terms $\kappa_1,\ldots,\kappa_n$ exemplify the relation denoted by the term Π^n.

So if we invoke Convention (17.2), the following would serve well as a definition of *logical existence* for individuals:

$$x\!\downarrow\; \equiv_{df}\; \exists F Fx$$

Any biconditional licensed by this definition is true. For consider any instance of the definition for a individual term κ that is substitutable for x. If κ has a denotation, then $\kappa\!\downarrow$ is intuitively true while $\exists F F\kappa$ is semantically true: in any model of classical second-order logic, there have to be at least two properties (one true of everything in the model and one true of nothing in the model) and so if κ has a denotation, the individual it denotes exemplifies a property. If κ fails to denote an individual, then $\kappa\!\downarrow$ is intuitively false and $\exists F F\kappa$ becomes false semantically: no formula of the form $\Pi\kappa$ is true and so the claim that κ exemplifies a property is false. So the biconditional licensed by the above definition is true no matter whether the term replacing x has a denotation or fails to have one.

However, if exemplification is the only mode of predication in the system, as it is in classical second-order logic, then one can't similarly say that a property *exists* if and only if some individual exemplifies it. This might be represented as the following definition:

$$F\!\downarrow\; \equiv_{df}\; \exists x Fx$$

The problem here is that F might be a property that provably isn't exemplified, such as *being concrete and failing to be concrete*, i.e., $[\lambda x\, E!x \,\&\, \neg E!x]$. In a theory on which $[\lambda x\, E!x \,\&\, \neg E!x]$ denotes a property, any biconditional licensed by the above definition, such as:

$$[\lambda x\, E!x \,\&\, \neg E!x]\!\downarrow\; \equiv \exists y([\lambda x\, E!x \,\&\, \neg E!x]y)$$

would be false. The left side would be true (since the λ-expression denotes a property), while the right side would be false (since nothing exemplifies the property in question, on pain of contradiction). This conclusion, of course, assumes that $[\lambda x E!x \& \neg E!x]$ obeys the principle $\forall y([\lambda x E!x \& \neg E!x]y \equiv E!y \& \neg E!y)$.

These observations may explain why some free logicians working within a classical exemplification logic intoduce a *primitive* notion of identity and define existence in terms of identity. Given a primitive notion of identity, the formula $\exists\beta(\beta = \tau)$, where β is any variable and τ is any term of the same type as β in which β doesn't occur free, says that there exists something identical to τ, i.e., there exists such a thing as τ. Consequently, one might find the following definition schema being used to define existence for *either* individual terms or relation terms:

$$\tau{\downarrow} \equiv_{df} \exists\beta(\beta = \tau), \text{ provided } \beta \text{ is a variable that doesn't occur free in } \tau$$

When τ is an individual term, such as $\imath xKx$, then the definition licenses the equivalence $\imath xKx{\downarrow} \equiv \exists y(y = \imath xKx)$. Intuitively, this biconditional is true: if $\imath xKx$ has a denotation, then both sides of the biconditional are true, and if $\imath xKx$ doesn't have a denotation, both sides of the biconditional are false. And similarly in the case of property terms, for the above definition also licenses the equivalence $[\lambda x\,\varphi]{\downarrow} \equiv \exists F(F = [\lambda x\,\varphi])$. If $[\lambda x\,\varphi]$ denotes a property (even one that nothing exemplifies), then both sides of the biconditional are true, and if it doesn't denote a property, then both sides are false.

Though the above definition of existence in terms of identity will *not* be adopted in the present system, it does explain why either the claim $\exists\beta(\beta = \tau)$ or the equivalent claim $\tau{\downarrow}$ would typically be used in the axioms for a negative free logic of quantification that govern a second-order language (with identity) that includes non-denoting (individual and relation) terms. In such logics, it is important to avoid applying, to non-denoting terms, rules of reasoning that are valid only for denoting terms. Using a mixture of material and formal mode, then the three main axioms of such a negative free logic assert:

- If everything is such that φ, then if τ has a *denotation*, then τ is such that φ.
- Primitive constants, variables, and certain *safe* λ-expressions (see below) have denotations.
- If an atomic formula is true, then the primary terms (7.8) in that formula have denotations.

Of course, when the above axioms are formulated in a second-order language, one uses either $\exists\beta(\beta = \tau)$ or $\tau{\downarrow}$ in place of the metalinguistic claim that τ has a denotation.

In fact, versions of the above axioms are preserved in the present theory and are asserted in the next chapter as axioms (39.1), (39.2), and (39.5.a) and (39.5.b). However, the key point here is that in systems with only one mode of predication, *identity* (=) is taken as a primitive, *existence* ($\downarrow$) is defined in terms of identity, and then either the claim $\tau{\downarrow}$ or the claim $\exists\beta(\beta = \tau)$ is used to state the axioms of the negative free logic.[106]

[106] I'm discounting the suggestion that one define identity for properties as either $F = G \equiv \forall x(Fx \equiv Gx)$ or $F = G \equiv \Box\forall x(Fx \equiv Gx)$, since such definitions collapse materially equivalent or necessarily equivalent properties.

By contrast, the present system has greater expressive and analytical power than the standard second-order predicate calculus. Our system is grounded on the additional principle:

> An encoding formula of the form $\kappa_1 \ldots \kappa_n \Pi^n$ ($n \geq 1$) is true if and only if the terms $\kappa_1, \ldots, \kappa_n$, and Π^n all have a denotation and the individuals denoted by the terms $\kappa_1, \ldots, \kappa_n$ encode the relation denoted by the term Π^n.

With two modes of predication, one can define *existence* (i.e., define $\tau\downarrow$) in terms of predication, as long as one uses the appropriate form of predication for the various cases. The keys to these definitions are the following intuitions, where κ is any individual term and Π any unary relation term: (a) $\ulcorner\kappa$ *exists*$\urcorner$ is true if and only if κ exemplifies some property, and (b) $\ulcorner\Pi$ *exists*$\urcorner$ is true if and only if some individual encodes Π. Definition (b) can then be generalized for n-ary relations of any arity. These definitions will be formally presented in (20.1) – (20.3), respectively. In Remark (33), we describe why (b) is justified and constitutes a good definition of property existence.

In light of these developments, we won't need identity as a primitive in order to define existence. Indeed, formulas of the form $\tau = \sigma$ can be defined and, in the case of n-ary relations ($n \geq 0$), will be defined in terms of *existence*. Then, once our deductive system is in place, *we shall prove as a theorem* (121.1) that the claim $\tau\downarrow \equiv \exists\beta(\beta = \tau)$ governs the *defined* notions $\tau\downarrow$ and $\tau = \sigma$. So, the axioms of our negative free logic of quantification will use the claim $\tau\downarrow$, as defined in our language, to capture the formal mode claim that τ has a denotation. The three axioms informally-stated in the three bullet points above will be formally represented by the following four axioms (where the third axiom stated informally above has been formalized using two separate claims):

- $\forall\alpha\varphi \to (\tau\downarrow \to \varphi^\tau_\alpha)$ (39.1)
- $\tau\downarrow$, provided τ is a primitive constant, a variable, or a core λ-expression (9.2) in which no variable bound by the λ occurs in encoding position (9.1) in the matrix (39.2)
- $\Pi^n\kappa_1 \ldots \kappa_n \to (\Pi^n\downarrow \,\&\, \kappa_1\downarrow \,\&\, \ldots \,\&\, \kappa_n\downarrow)$ $(n \geq 0)$, where $\Pi^n\kappa_1 \ldots \kappa_n$ is any exemplification formula (39.5.a)
- $\kappa_1 \ldots \kappa_n\Pi^n \to (\Pi^n\downarrow \,\&\, \kappa_1\downarrow \,\&\, \ldots \,\&\, \kappa_n\downarrow)$ $(n \geq 1)$, where $\kappa_1 \ldots \kappa_n\Pi^n$ is any encoding formula (39.5.b)

Consequently, *primitive* notions of existence and identity won't be needed to state the above axioms.[107] The above axioms therefore have greater significance than their counterparts in languages where $\downarrow$ or $=$ is primitive. The definitions in (20) and in (23.1) – (23.4) reveal that the two notions of predication

[107] Cf. Feferman 1995 (299), which includes primitive atomic formulas of the form $s = t$, $t\downarrow$, and $\mathbf{R}(t_1, \ldots, t_n)$. Compare also his 'Definedness' axioms on p. 301.

that serve as the basis of object theory are more fundamental than the notions of existence and identity, and may thereby offer insightful definitions of both notions.

(**31**) **Remark:** Explanation of the Substitutability Requirement in Convention (17.2.a). To illustrate the substitutability requirement in Convention (17.2.a), consider the definition of $x{\downarrow}$ in (20.1), i.e.,

$$x{\downarrow} \equiv_{df} \exists F Fx \qquad (20.1)$$

By Convention (17.2.a), we can uniformly replace x by any individual term κ to obtain an instance of this definition, *provided* κ is substitutable for x in $\exists F Fx$. So the Convention says that the following are instances of the definition, since y and $\imath zGz$ are substitutable for x in the definiens:

$$(x{\downarrow})^{y}_{x} \equiv_{df} (\exists F Fx)^{y}_{x}, \text{ i.e.,}$$
$$y{\downarrow} \equiv_{df} \exists F Fy$$

$$(x{\downarrow})^{\imath zGz}_{x} \equiv_{df} (\exists F Fx)^{\imath zGz}_{x}, \text{ i.e.,}$$
$$\imath zGz{\downarrow} \equiv_{df} \exists F F\imath zGz$$

However, the substitutability requirement rules out the following as an instance of (20.1):

$$(\omega)\quad (x{\downarrow})^{\imath zFz}_{x} \equiv_{df} (\exists F Fx)^{\imath zFz}_{x}, \text{ i.e.,}$$
$$\imath zFz{\downarrow} \equiv_{df} \exists F F\imath zFz$$

This is ruled out because $\imath zFz$ is not substitutable for x in the definiens of (20.1), as required by Convention (17.2.a).

In other words, we're understanding (20.1) as the following schema, in which the individual variable x is replaced by the metavariable κ:

(δ) $\kappa{\downarrow} \equiv_{df} \exists F F\kappa$, provided F doesn't occur free in κ.

The proviso on (δ) has to be added so that only terms substitutable for x in the definiens of (20.1) can be used to obtain instances of the definition. This illustrates footnote 86 to Convention (17.2.a), which tells us:

> each distinct object-language variable α that occurs free in a definition functions as a distinct Greek metavariable that ranges over terms of the same type as α, with the proviso that if α falls within the scope of a variable-binding operator **Op** in the definiens, then a term may serve as the instance of α only if it doesn't contain free occurrences of the variable bound by **Op**.

If x in (20.1) is functioning in the same way as κ does in (δ), the Convention requires that since x falls within the scope of the variable-binding operator $\forall F$

in the definiens, the terms that serve as instances of x may not contain free occurrences of F. (ω) would change the meaning of the definition and increase the burden of proof for the claim $\imath xFx\!\downarrow$. Intuitively, to prove $\imath zFz$ exists, it suffices to show that $\imath zFz$ exemplifies some property or other; it need not be F.

If Convention (17.2.a) rules out (ω) as an instance of (20.1), then we can't as yet obtain an instance of the definition for the term $\imath zFz$. So what is the definition of the expression $\imath zFz\!\downarrow$? This is the topic of the next Remark, where we discuss Convention (17.2.b).

(**32**) **Remark:** Discussion of Alphabetically-Variant Definientia. Consider the definition of $x\!\downarrow$ in (20.1), i.e.,

$$x\!\downarrow \equiv_{df} \exists F Fx \tag{20.1}$$

We saw in Remark (28) why the variable F in (20.1) has to function as a metavariable, so that the following become instances of the definition:

$$x\!\downarrow \equiv_{df} \exists G\, Gx$$

$$x\!\downarrow \equiv_{df} \exists H\, Hx$$

Moreover, we saw in Remark (31) that we cannot substitute every individual term for x in (20.1) to produce an instance of the definition. In particular, we saw that we cannot obtain the instance:

(ζ) $\imath zFz\!\downarrow \equiv_{df} \exists F\, F\imath zFz$

But the following would be a viable definition of $\imath zFz\!\downarrow$:

(ϑ) $\imath zFz\!\downarrow \equiv_{df} \exists G\, G\imath zFz$

To see how Convention (17.2) implies that (ϑ) is an instance of (20.1), note that by the (17.2.a), the following is an instance:

$$y\!\downarrow \equiv_{df} \exists F\, Fy$$

So by Convention (17.2.b), the following is an instance, since the definiens is an alphabetic-variant of the definiens in the above:

$$y\!\downarrow \equiv_{df} \exists G\, Gy$$

But, Convention (17.2.a) tells us that the free occurrence of G in the above functions as a metavariable, and that the following is therefore also an instance:

$(y\!\downarrow)^{\imath zFz}_{y} \equiv_{df} (\exists G\, Gy)^{\imath zFz}_{y}$, i.e.,
$\imath zFz\!\downarrow \equiv_{df} \exists G\, G\imath zFz$

By this means, (ϑ) becomes a well-formed instance of (20.1) and we obtain a well-formed definition of $\imath zFz\downarrow$.

Thus, given the two clauses in Convention (17.2), (20.1) becomes equivalent to the following definition schema, in which Ω is a metavariable ranging over unary relation variables:

(ξ) $\kappa\downarrow \equiv_{df} \exists\Omega\,\Omega\kappa$,
provided Ω doesn't occur free in κ

This understanding of definition (20.1) illustrates footnote 87 to Convention (17.2.b), which tells us:

> each distinct object-language variable α that occurs bound by a variable-binding operator **Op** in the definiens functions as a distinct Greek meta-variable ranging over variables of the same type as α, with the proviso that an object-language variable may serve as an instance of α only if it does not occur free in any term occurring within the scope of **Op**.

Since F in (20.1) is functioning as Ω does in the (ξ), we can't obtain an instance of the definition by replacing F with some a variable that occurs free within a term substituted for κ, since κ falls within the scope of $\exists F$. The reader should now convince him- or herself that every instance of the following is an instance of (20.1), for any individual variable ν:

$\imath\nu\varphi\downarrow \equiv_{df} \exists F\, F\imath\nu\varphi$,
provided F doesn't occur free in φ

Finally, once it is established that alphabetically-variant formulas ψ and ψ' are equivalent and interderivable, then we shall be able to prove that the definitions $\varphi \equiv_{df} \psi$ and $\varphi \equiv_{df} \psi'$ are equivalent. But, in the present system, the proof that alphabetically-variant formulas are equivalent and interderivable doesn't occur until (111.5) and (114), respectively. So, until that time, we must ensure that definitions-by-$\equiv$ have instances in which the definientia are alphabetic variants, so that we can produce instances of the definition for every open term that might be substituted for the free variables, as in the case of (20.1) and (ϑ) above.

Now, to see how an alphabetically-variant definiens is needed for a definition-by-=, consider the following definition of property negation:

(ω) $\overline{F} =_{df} [\lambda x\, \neg Fx]$

Since F functions as a metavariable ranging over unary relation terms, consider the term $[\lambda y\, Rxy]$. Note that $[\lambda y\, Rxy]$ is not substitutable for F in the definiens of (ω). So the following is *not* an instance of (ω):

$\overline{[\lambda y\, Rxy]} =_{df} [\lambda x\, \neg[\lambda y\, Rxy]x]$

Consequently, to produce a definition of $\overline{[\lambda y\, Rxy]}$, we note that, by Convention (17.2.b), the following is an instance of (ω), since its definiens is an alphabetic variant of the definiens in (ω):

$$\overline{F} =_{df} [\lambda z\, \neg Fz]$$

But then, the variable F in the above is also functioning as a metavariable, and so by Convention (17.2.a), the following is an instance:

$$\overline{[\lambda y\, Rxy]} =_{df} [\lambda z\, \neg[\lambda y\, Rxy]z]$$

This illustrates how Convention (17.2.b) applies in definitions-by-=. As an exercise, the reader should explain how this example also illustrates footnote 87.

(**33**) **Remark:** Justification of the Definition of n-ary Relation Existence For $n \geq 1$. It would serve well to explain in some detail why definition (20.2) is justified. Consider first the unary case of the definition, i.e., $F{\downarrow} \equiv_{df} \exists x\, xF$. To see that this is a good definition, we need only show that if $\downarrow$ had been an undefined, primitive operator yielding formulas of the form $\Pi{\downarrow}$, then it would have been a theorem that $F{\downarrow} \equiv \exists x\, xF$.

To see that this would have been a theorem, we'll again need to appeal to a number of facts not yet in evidence. In Remark (248) we note that our theory implies, without an appeal to the definition of existence, that every property is encoded by some object, i.e., $\forall H \exists x\, xH$. This result is a relatively straightforward consequence of the Comprehension Principle for Abstract Objects, which is asserted as an axiom in the next chapter, in item (53).[108]

With this fact, it is easy to see why $F{\downarrow} \equiv \exists x\, xF$ would have been a theorem if $\downarrow$ had been primitive. For the left-to-right direction, assume $F{\downarrow}$. Then by axiom (39.1) of our free quantificational logic for complex terms and the fact that F doesn't occur free in $\exists x\, xH$ we can instantiate F into $\forall H \exists x\, xH$ to conclude $\exists x\, xF$. For the right-to-left direction, assume $\exists x\, xF$. Assume further a is a witness to this claim, so that we know aF. Then, by axiom (39.5.b) of our free quantificational logic for complex terms, it follows that $F{\downarrow} \,\&\, a{\downarrow}$. So by Rule &E (86.2.a), $F{\downarrow}$.

Since it would have been a theorem that $F{\downarrow} \equiv \exists x\, xF$ if $\downarrow$ had been a primitive notion instead of a defined one, our system would have provided us with an *analysis* of property existence. Analogous reasoning shows that, for

[108] To anticipate the discussion in Remark (248), let H be an arbitrary property and note that the Comprehension Principle for Abstract Objects has the following instance:

$$\exists x(A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Hz))$$

So where a is an arbitrary such object, we know $\forall F(aF \equiv \forall z(Fz \equiv Hz))$. Since we know $H{\downarrow}$ by axiom (39.2) and H is substitutable for F in the matrix $aF \equiv \forall z(Fz \equiv Hz)$, instantiate to H so that we can conclude $aH \equiv \forall z(Hz \equiv Hz)$. But $\forall z(Hz \equiv Hz)$ is a simple theorem of predicate logic. Hence, it follows that aH, and further follows that $\exists x\, xH$. Since H was arbitarily chosen, we may conclude $\forall H \exists x\, xH$.

$n > 1$, if $\downarrow$ had been primitive, it would have been a theorem that $F^n\downarrow \equiv \exists x_1 \ldots \exists x_n(x_1 \ldots x_n F^n)$. In this case, the result is grounded in the provability of $\forall G^n \exists x_1 \ldots \exists x_n(x_1 \ldots x_n G^n)$ *without* an appeal to the definition of existence. This is also discussed in item (248).[109] Since these would have been facts, we reduced the number of primitives by taking the analysis of relation existence to be a definition. Our system is powerful enough to define relation existence in terms of predication.

(**34**) **Remark:** Digression on the Definition of Relation Existence. It may be of interest to see why one can't define relation existence in terms of property existence. Suppose one were to define, for $n \geq 2$:[110]

$$(\vartheta)\ \ F^n\downarrow \equiv_{df} \forall y_1 \ldots \forall y_{n-1}([\lambda x\, F^n x y_1 \ldots y_{n-1}]\downarrow \,\&\, [\lambda x\, F^n y_1 x y_2 \ldots y_{n-1}]\downarrow \,\&\, \ldots \,\&\, [\lambda x\, F^n y_1 \ldots y_{n-1} x]\downarrow)$$

(ϑ) stipulates that F^n *exists* just in case all of the properties projectable from it exist, i.e., just in case all of the properties that result from the different ways of plugging $n-1$ objects into F^n exist. The problem with (ϑ) is that it would implicitly introduce, as axioms, biconditionals that are provably false.

In particular, our conventions for definitions would introduce the following biconditional as axiomatic, where Π is any binary relation term in which x and y don't occur free:

$$(\xi)\ \ \Pi\downarrow \equiv \forall y([\lambda x\, \Pi x y]\downarrow \,\&\, [\lambda x\, \Pi y x]\downarrow)$$

But to see that (ξ) is false, we need only show that the right condition of (ξ) is true when the left side is false. So suppose $\neg\Pi\downarrow$. To show that the right side holds, we need only show $[\lambda x\, \Pi x y]\downarrow \,\&\, [\lambda x\, \Pi y x]\downarrow$, since y doesn't occur free in our assumption. But both conjuncts will be instances of axiom (39.2); by hypothesis, x doesn't occur free in Π, and so the λ-expressions in each conjunct are core (9.2) – no variable bound by the λ occurs in encoding position (9.1) in the matrix. So, the right side of (ξ) is true, while the left side is false. Hence, one can't define relation existence in terms of property existence as in (ϑ).

---

[109] To preview the discussion, the simplest proof of this claim relies on two facts, namely, axiom (50), which asserts that n-ary encoding facts of the form $x_1 \ldots x_n G$ are equivalent to a conjunction of unary encodings of the properties that can be projected from G, and the fact that there is an abstract object, say a, that encodes every property. Note first that axiom (39.2) guarantees that all of the following properties exist: $[\lambda x\, Gxa \ldots a]$, $[\lambda x\, Gaxa \ldots a]$, $\ldots$, and $[\lambda x\, Ga \ldots ax]$. Since $\forall F aF$, a encodes all of these latter properties. So we can conclude:

$$[\lambda x\, Gxa \ldots a] \,\&\, a[\lambda x\, Gaxa \ldots a] \,\&\, \ldots \,\&\, a[\lambda x\, Ga \ldots ax]$$

Hence, by axiom (50), $a \ldots aG$, and so $\exists x_1 \ldots \exists x_n(x_1 \ldots x_n G)$.

[110] This was observed by Daniel Kirchner (personal communication, 19 September 2017), though the reason for the conclusion has subsequently changed from his original, given subsequent improvements made to axiom (39.2).

(**35**) **Remark:** About Identity for Individuals. It is important to understand the behavior of definition (23.1) with respect to definite descriptions. Consider a description such as $\imath zGz$. Then we obtain instances of (23.1) in which $\imath zGz$ uniformly replaces x or y. So, for example, the following becomes an instance of the definition:

$$(\vartheta)\quad x = \imath zGz \equiv_{df} (O!x \,\&\, O!\imath zGz \,\&\, \Box\forall F(Fx \equiv F\imath zGz)) \vee (A!x \,\&\, A!\imath zGz \,\&\, \Box\forall F(xF \equiv \imath zGzF))$$

It is important to observe:

(a) The definiens of (ϑ) implies $\imath zGz\downarrow$, since both disjuncts imply it. The left disjunct implies $O!\imath zGz$, which in turn implies $\imath zGz\downarrow$, by axiom (39.5.a). This axiom guarantees that the primary terms (7.8) of an exemplification or encoding formula are significant; this will be discussed in a bit more detail in (231). Moreover, the right disjunct implies $A!\imath zGz$, which also implies $\imath zGz\downarrow$, again by axiom (39.5.a).

(b) Consequently, if it is known, either by hypothesis or by proof, that $\neg\imath zGz\downarrow$, then the negation of the definiens becomes derivable. So if (ϑ) implicitly introduces a biconditional axiom in which $\equiv$ replaces $\equiv_{df}$, then the negated identity $\neg(x = \imath zGz)$ becomes derivable. Since this holds for any x, the claim $\forall x\neg(x = \imath zGz)$ becomes derivable for any description $\imath zGz$ for which it is known that $\neg\imath zGz\downarrow$.

These facts play an important role as we develop and apply our system.

(**36**) **Remark:** Why Existence Clauses Are Sometimes Needed in the Definiens of Definitions-by-$\equiv$. Before we consider specifically why existence clauses have been added to the definiens of (23.2), we discuss the more general problem that arises for definitions-by-$\equiv$. To see the problem, suppose one wanted to say that a property *is conditionally necessary for* an object just in case the object necessarily exemplifies the property whenever it exemplifies the property. Then we would formalize this definition as:

$$CondNecFor(F,x) \;\equiv_{df}\; Fx \to \Box Fx$$

with the understanding that, by Convention (17.2), the above is to be interpreted as the following definition schema:

$$(A)\quad CondNecFor(\Pi,\kappa) \;\equiv_{df}\; \Pi\kappa \to \Box\Pi\kappa$$

Suppose further that this schema introduces the instances (and their closures) of the following schema as axioms:

$$(B)\quad CondNecFor(\Pi,\kappa) \equiv (\Pi\kappa \to \Box\Pi\kappa)$$

Now consider an arbitrary description, say $\imath z\psi$, and an arbitrary unary λ-expression, say $[\lambda x\,\varphi]$, and suppose that both $\neg\imath z\psi\downarrow$ and $\neg[\lambda x\,\varphi]\downarrow$ are known, either by hypothesis or by proof. Then where P and a are constants (for which it will be axiomatic that $P\downarrow$ and $a\downarrow$), then we *should* be able to derive the following from (B):

(C) $\neg \mathit{CondNecFor}(P, \imath z\psi)$

(D) $\neg \mathit{CondNecFor}([\lambda x\,\varphi], a)$

(E) $\neg \mathit{CondNecFor}([\lambda x\,\varphi], \imath z\psi)$

In general, we should be able to derive $\neg \mathit{CondNecFor}(\Pi, \kappa)$ whenever it is known that either Π or κ fails to denote. Such a derivation would preserve the *garbage in, garbage out* principle. However, one cannot derive such claims from (B). In fact, one can derive:

(F) $\mathit{CondNecFor}(P, \imath z\psi)$

(G) $\mathit{CondNecFor}([\lambda x\,\varphi], a)$

(H) $\mathit{CondNecFor}([\lambda x\,\varphi], \imath z\psi)$

Specifically, from $\neg\imath z\psi\downarrow$ and (B), one can derive (F) and (H), and from $\neg[\lambda x\,\varphi]\downarrow$ and (B), one can derive (G) and (H). Here's how.

To see that (F) and (H) are derivable from (B) and $\neg\imath z\psi\downarrow$, note that our system will allow one to prove, from the latter, the negation of any exemplification or encoding formula in which $\imath z\psi$ occurs as a primary term. This is guaranteed by axioms (39.5.a) and (39.5.b); they are grounded in the fact that an n-ary predication ($n \geq 1$) is true if and only if its primary terms have denotations and the objects denoted by the individual terms exemplify or encode (as the case may be) the relation denoted by the relation term. Consequently, the negations of $P\imath z\psi$ and $[\lambda x\,\varphi]\imath z\psi$ both follow from $\neg\imath z\psi\downarrow$. But then, the conditionals $P\imath z\psi \rightarrow \Box P\imath z\psi$ and $[\lambda x\,\varphi]\imath z\psi \rightarrow \Box[\lambda x\,\varphi]\imath z\psi$ are both derivable, by failure of the antecedent. Given (B), these respectively imply $\mathit{CondNecFor}(P, \imath z\psi)$, i.e., (F), and $\mathit{CondNecFor}([\lambda x\,\varphi], \imath z\psi)$, i.e., (H). Since (B) is implicitly introduced by (A), the latter leads to violations of the *garbage in, garbage out* principle; given (A), *CondNecFor* becomes a condition that *holds* even though its second argument fails to denote. So we have instances of the principle *garbage in, roses out*, something one should like to avoid when one can control it. We should be able to prove (C) and (E) when $\imath z\psi$ fails to denote.

By analogous reasoning, (G) and (H) are derivable from (B) and $\neg[\lambda x\,\varphi]\downarrow$. The latter implies the negations of $[\lambda x\,\varphi]a$ and $[\lambda x\,\varphi]\imath z\psi$. So the conditionals $[\lambda x\varphi]a \rightarrow \Box[\lambda x\varphi]a$ and $[\lambda x\varphi]\imath z\psi \rightarrow \Box[\lambda x\varphi]\imath z\psi$ are both derivable by failure of the antecedent. Given (B), these respectively imply that $\mathit{CondNecFor}([\lambda x\,\varphi], a)$, i.e., (G), and $CondNecFor([\lambda x\, \varphi]b, \imath z\psi)$, i.e., (H). Since (B) is implicitly introduced by (A), the latter again leads to violations of the *garbage in, garbage out* principle; given (A), *CondNecFor* becomes a condition that *holds* even though its first argument fails to denote. We have further instances of the principle *garbage in, roses out*. We should be able to derive (D) and (E) when it is known, either by hypothesis or proof, that $[\lambda x\, \varphi]$ isn't significant.

To solve this problem in a general way, we need to add existence conditions to the definiens. In our example, we can reformulate our original definition, under Convention (17.2), as:

$$CondNecFor(F, x) \equiv_{df} F\downarrow \,\&\, x\downarrow \,\&\, (Fx \rightarrow \Box Fx)$$

where this is to be interpreted as the schematic definition:

(A′) $CondNecFor(\Pi, \kappa) \equiv_{df} \Pi\downarrow \,\&\, \kappa\downarrow \,\&\, (\Pi\kappa \rightarrow \Box\Pi\kappa)$

(A′) would implicitly extend our system with all of the instances, and the closures of the instances, of the following axiom or theorem schema:

(I) $CondNecFor(\Pi, \kappa) \equiv (\Pi\downarrow \,\&\, \kappa\downarrow \,\&\, (\Pi\kappa \rightarrow \Box\Pi\kappa))$

Thus, the following would all be theorems, by uniformly assigning terms to metavariables in (I):

(J) $CondNecFor(P, \imath z\psi) \equiv (P\downarrow \,\&\, \imath z\psi\downarrow \,\&\, (P\imath z\psi \rightarrow \Box P\imath z\psi))$

(K) $CondNecFor([\lambda x\, \varphi], a) \equiv ([\lambda x\, \varphi]\downarrow \,\&\, a\downarrow \,\&\, ([\lambda x\, \varphi]a \rightarrow \Box[\lambda x\, \varphi]a))$

(L) $CondNecFor([\lambda x\, \varphi], \imath z\psi) \equiv ([\lambda x\, \varphi]\downarrow \,\&\, \imath z\psi\downarrow \,\&\, ([\lambda x\, \varphi]\imath z\psi \rightarrow \Box[\lambda x\, \varphi]\imath z\psi))$

So from $\neg\imath z\psi\downarrow$ and (J) we can infer (C), and from $\neg\imath z\psi\downarrow$ and (L) we can infer (E). Analogously, from $\neg[\lambda x\, \varphi]\downarrow$ and (K) we can infer (D), and from $\neg[\lambda x\, \varphi]\downarrow$ and (L) we can infer (E). This is as it should be: given (A′), the negation of $CondNecFor(\Pi, \kappa)$ is derivable whenever $\neg\Pi\downarrow$ or $\neg\kappa\downarrow$ is known either by proof or by hypothesis. This gives us a way to preserve the *garbage in, garbage out* principle.

We're not suggesting here that every definition-by-$\equiv$ with free variables needs to include existence clauses in the definiens. There are a number of cases where they aren't needed. In some cases, existence is already implied by the definiens, such as in the example from Remark (27):

$$ContingentlyExemplifies(x, F) \equiv_{df} Fx \,\&\, \neg\Box Fx$$

This definition doesn't require existence clauses in the definiens, since no matter what terms Π and κ uniformly replace F and x, respectively, in the definition, the definiens implies both $\Pi\downarrow$ and $\kappa\downarrow$. In other cases, we don't specially need or want the definiens to imply existence, such as in the definitions of &,

$\vee$, $\equiv$, $\exists\alpha\varphi$, and $\Diamond\varphi$ in (18), or in definition of $\tau \neq \sigma$ in (24). And, there will be some cases where, in order to obtain a correct definiens, existence clauses have to be added to the definiens for some but not all of the variables that occur free in the definition.

Now that we have a good grasp on why existence clauses are sometimes needed in the definiens of definitions-by-$\equiv$, we can return specifically to definition (23.2) of $F = G$. In this case, the existence clauses in the definiens are *essential*. To see why, suppose we had omitted them and consider the following instance of the definition *without* the existence clauses in the definiens, in which $[\lambda y\, \varphi]$ uniformly replaces F and $[\lambda z\, \psi]$ uniformly replaces G:

(ϑ) $[\lambda y\, \varphi] = [\lambda z\, \psi] \ \equiv_{df}\ \Box\forall x(x[\lambda y\, \varphi] \equiv x[\lambda z\, \psi])$

Now suppose that this definition were to introduce the following as an axiom:

(ϑ') $[\lambda y\, \varphi] = [\lambda z\, \psi] \equiv \Box\forall x(x[\lambda y\, \varphi] \equiv x[\lambda z\, \psi])$

Finally, suppose both $[\lambda y\, \varphi]$ and $[\lambda z\, \psi]$ are both λ-expressions that fail to have a denotation, i.e., suppose $\neg[\lambda y\, \varphi]\downarrow$ and $\neg[\lambda z\, \psi]\downarrow$. Then ($\vartheta'$) would yield a derivation of $[\lambda y\, \varphi] = [\lambda z\, \psi]$. Here's why.

As mentioned previously, our negative free logic of quantification for complex terms will require, by axioms asserted in the next chapter, (39.5.a) and (39.5.b), that if an n-ary exemplification or encoding formula ($n \geq 1$) is true, then all the primary terms in that formula are significant. So, by the contrapositive, it follows from $\neg[\lambda y\, \varphi]\downarrow$ and $\neg[\lambda z\, \psi]\downarrow$ that the encoding formulas $x[\lambda y\, \varphi]$ and $x[\lambda z\, \psi]$ are both false. Since they both have the same truth value, it follows by propositional logic that $x[\lambda y\, \varphi] \equiv x[\lambda z\, \psi]$. Since this holds for arbitrary x, it follows by quantificational logic that $\forall x(x[\lambda y\, \varphi] \equiv x[\lambda z\, \psi])$. Since this universal claim was established without appealing to any contingent assumptions, it follows by modal logic that $\Box\forall x(x[\lambda y\, \varphi] \equiv x[\lambda z\, \psi])$. But if ($\vartheta$) implicitly introduces (ϑ'), then we would have a proof of $[\lambda y\, \varphi] = [\lambda z\, \psi]$. Clearly, we don't want to prove that an identity holds between *every* pair of non-denoting λ-expressions! Indeed, we don't want any pair of non-denoting λ-expressions Π and Π' to be such that $\Pi = \Pi'$.

We've adopted the solution of adding the clauses $F\downarrow$ and $G\downarrow$ in (23.2). This avoids the derivation of identities in which the terms flanking the identity sign fail to denote. It instead implies that if a property identity statement is true, then the terms flanking the identity sign have a denotation. Consequently, if Π is any property term that fails to be significant, so that $\Pi\downarrow$ is known to be false, the negation of the definiens of (23.2) will be derivable, no matter whether Π replaces F or G in the definition. So we'll be able to infer $\neg(\Pi = \Pi')$ if either $\neg\Pi\downarrow$ or $\neg\Pi'\downarrow$.

(**37**) **Remark:** Digression on n-ary Relation Identity. Some readers may wonder why we haven't defined relation identity by using n-ary encoding to generalize the definition of property identity, i.e., as follows, where F and G are n-ary relation variables ($n \geq 2$):

$$(\vartheta)\;\; F = G \;\equiv_{df}\; F{\downarrow} \;\&\; G{\downarrow} \;\&\; \Box\forall x_1 \ldots \forall x_n(x_1 \ldots x_n F \equiv x_1 \ldots x_n G)$$

The problem with (ϑ) is that it implies: if there is exactly one necessarily unexemplified property, then there is at most one ordinary object.[111] Here is the argument in full, though it may be skipped and reserved for later study without loss of understanding of what follows. However, the the argument makes use of a number of concepts and theorems that aren't derived until later. For example, we'll reference the relation constant $=_E$, which isn't introduced and proved to be significant until (230), and reference theorems about $=_E$, such as its reflexivity on the ordinary objects (239.1).

We first derive some preliminary consequences of (ϑ). Suppose that the inferential role of (ϑ) is to implicitly introduce (the closures of) its instances as axioms. Then, if (ϑ) had been offered as a definition, the following special case for binary relations would become axiomatic, where F and G are object-language variables:

$$(\vartheta')\;\; F = G \;\equiv\; (F{\downarrow} \;\&\; G{\downarrow} \;\&\; \Box\forall x\forall y(xyF \equiv xyG))$$

Since $F{\downarrow}$ and $G{\downarrow}$ are axiomatic (39.2), (ϑ') reduces, by Rule $\equiv$S (91), to:

$$F = G \equiv \Box\forall x\forall y(xyF \equiv xyG)$$

Though it may not be immediately apparent why, this is equivalent to:

$$(\zeta)\;\; F = G \equiv \forall x\forall y(xyF \equiv xyG)$$

The reason they are equivalent is grounded in the axiom for the rigidity of encoding (51), i.e., that $xF \rightarrow \Box xF$, and we leave a full demonstration to a footnote.[112] Now the axiom governing n-ary encoding (50) has the following instances:

[111] This was observed by Daniel Kirchner, who outlined the argument in personal communications (computer audio discussions and email exchanges) of 9–11 November 2017.

[112] The reasoning that shows that $\forall x\forall y(xyF \equiv xyG)$ is equivalent to $\Box\forall x\forall y(xyF \equiv xyG)$ depends on the Barcan Formula and the binary case of theorem (179.6), which is derived from the axiom for the rigidity of encoding (51) and the theorem (178.1) that is derivable from it. The binary case of (179.6) is:

$$(x_1x_2F \equiv y_1y_2G) \equiv \Box(x_1x_2F \equiv y_1y_2G)$$

(The derivation of this fact from (51) and (178.1) is in the Appendix.) By 4 applications of GEN:

$$\forall x_1\forall x_2\forall y_1\forall y_2((x_1x_2F \equiv y_1y_2G) \equiv \Box(x_1x_2F \equiv y_1y_2G))$$

Instantiating the quantifiers, in order, to x, y, x, and y, respectively, it follows that:

$$(xyF \equiv xyG) \equiv \Box(xyF \equiv xyG)$$

So by GEN:

$$\forall x\forall y((xyF \equiv xyG) \equiv \Box(xyF \equiv xyG))$$

$$xyF \equiv (x[\lambda z\, Fzy] \,\&\, y[\lambda z\, Fxz])$$

$$xyG \equiv (x[\lambda z\, Gzy] \,\&\, y[\lambda z\, Gxz])$$

Since these are equivalences that don't depend on a contingency, we may use the Rule of Substitution to substitute, in (ζ), the right condition of these equivalences for the left, to conclude:

$$(\xi)\quad F = G \equiv \forall x \forall y((x[\lambda z\, Fzy] \,\&\, y[\lambda z\, Fxz]) \equiv (x[\lambda z\, Gzy] \,\&\, y[\lambda z\, Gxz]))$$

But notice, independently, that it is axiomatic (52) that ordinary objects fail to encode properties. So whenever x or y is an *ordinary* object, the (quantified) biconditional in the right condition of (ξ), i.e., $(x[\lambda z\, Fzy] \,\&\, y[\lambda z\, Fxz]) \equiv (x[\lambda z\, Gzy] \,\&\, y[\lambda z\, Gxz])$ will be true — the conjunctions on both sides of the biconditional are false whenever x or y is ordinary, since in either case, one of the conjuncts (all of which are encoding formulas) will be false. In other words, it is provable without appealing to any contingent assumptions (exercise) that:

Fact: $\forall x \forall y(\neg(A!x \,\&\, A!y) \rightarrow (x[\lambda z\, Fzy] \,\&\, y[\lambda z\, Fxz]) \equiv (x[\lambda z\, Gzy] \,\&\, y[\lambda z\, Gxz]))$

Given this Fact (which is a necessary truth given that no contingent assumptions are needed to prove it), (ξ) implies:

$$(\xi')\quad F = G \equiv \forall x \forall y((A!x \,\&\, A!y) \rightarrow (x[\lambda z\, Fzy] \,\&\, y[\lambda z\, Fxz]) \equiv (x[\lambda z\, Gzy] \,\&\, y[\lambda z\, Gxz]))$$

A strict proof that (ξ) implies (ξ') is left to a footnote.[113] Note that (ξ') holds for any relations F and G.

Now by two applications of the basic quantifier logic theorem $\forall\alpha(\varphi \equiv \psi) \equiv (\forall\alpha\varphi \equiv \forall\alpha\psi)$ (99.3), it follows that:

$$\forall x \forall y(xyF \equiv xyG) \equiv \forall x \forall y \Box(xyF \equiv xyG)$$

But the right side of this last biconditional is equivalent, by two instances of the Barcan formula (167.1) and a Rule of Substitution (160.2) to $\Box\forall x \forall y(xyF \equiv xyG)$. Hence:

$$\forall x \forall y(xyF \equiv xyG) \equiv \Box\forall x \forall y(xyF \equiv xyG) \qquad \bowtie$$

[113] We simplify the proof that (ξ) implies (ξ') by reasoning with respect to their form. Note that if we let ψ be the formula $(x[\lambda z\, Fzy] \,\&\, y[\lambda z\, Fxz]) \equiv (x[\lambda z\, Gzy] \,\&\, y[\lambda z\, Gxz])$, then ($\xi$) has the form:

$$(\xi)\quad F = G \equiv \forall x \forall y \psi$$

Now let χ be the formula $(A!x \,\&\, A!y)$, so that (ξ') has the form:

$$(\xi')\quad F = G \equiv \forall x \forall y(\chi \rightarrow \psi)$$

and the Fact has the form:

$$\forall x \forall y(\neg\chi \rightarrow \psi)$$

Then we can reason that (ξ) implies (ξ') as follows. Assume (ξ), i.e., $F = G \equiv \forall x \forall y \psi$. To show ($\xi'$), we establish both directions. ($\rightarrow$) Assume $F = G$. Then by (ξ), $\forall x \forall y \psi$. But this easily implies $\forall x \forall y(\chi \rightarrow \psi)$, since $\psi \rightarrow (\chi \rightarrow \psi)$ is the first law of propositional logic (38.1). ($\leftarrow$) Assume $\forall x \forall y(\chi \rightarrow \psi)$. To show $F = G$, it suffices by (ξ) to show $\forall x \forall y \psi$. By GEN, it suffices to show ψ. Assume $\neg\psi$ for reductio. But our local assumption, i.e., $\forall x \forall y(\chi \rightarrow \psi)$, implies $\chi \rightarrow \psi$. Hence, $\neg\chi$. But our Fact, namely $\forall x \forall y(\neg\chi \rightarrow \psi)$, implies $\neg\chi \rightarrow \psi$. Hence ψ. Contradiction.

Now recall that we are trying to show that the claim, there is a unique property that is necessarily unexemplified, i.e.,

$$\exists H(\Box\neg\exists xHx \,\&\, \forall H'(\Box\neg\exists xH'x \rightarrow H'=H))$$

implies that there is at most one ordinary object, i.e.,

$$\forall x\forall y(O!x \,\&\, O!y \rightarrow x=y)$$

So, for conditional proof, assume there is exactly one necessarily unexemplified property. To show that there is at most one ordinary object, it suffices to show $(O!a \,\&\, O!b) \rightarrow a=b$, where a and b are arbitrarily chosen. So assume $O!a$ and $O!b$. Then since $=_E$ is reflexive on the ordinary objects (239.1), it follows that $a=_E a$ and $b=_E b$. Now consider the relations:

(R_1) $[\lambda uv\; u=_E a \,\&\, v=_E a]$

(S_1) $[\lambda uv\; u=_E b \,\&\, v=_E b]$

Both provably exist.[114] So from the fact that $a=_E a$, it follows by strengthened β-Conversion (181) that R_1aa. Note that if we can show $R_1 = S_1$, then it follows that $a=b$ and we're done. For suppose $R_1 = S_1$. Then from the previously established R_1aa, it follows that S_1aa, and so $a=_E b \,\&\, a=_E b$, again by (181). *A fortiori*, $a=_E b$ and therefore $a=b$, by the definition of $=$ (23.1).

So it remains to show $R_1 = S_1$. Since (ξ') holds for any two (existing) relations, it suffices to show:

$$\forall x\forall y((A!x \& A!y) \rightarrow ((x[\lambda z R_1 zy] \& y[\lambda z R_1 xz]) \equiv (x[\lambda z S_1 zy] \& y[\lambda z S_1 xz])))$$

And so by GEN, it suffices to show:

$$(A!x \,\&\, A!y) \rightarrow ((x[\lambda z\, R_1 zy] \,\&\, y[\lambda z\, R_1 xz]) \equiv (x[\lambda z\, S_1 zy] \,\&\, y[\lambda z\, S_1 xz]))$$

So assume $A!x \& A!y$. Now if we can establish the identities $[\lambda z R_1 zy] = [\lambda z S_1 zy]$ and $[\lambda z\, R_1 xz] = [\lambda z\, S_1 xz]$, then we can establish $(x[\lambda z\, R_1 zy] \,\&\, y[\lambda z\, R_1 xz]) \equiv (x[\lambda z\, S_1 zy] \,\&\, y[\lambda z\, S_1 xz])$. For if we assume the left condition of this biconditional, $x[\lambda z R_1 zy] \& y[\lambda z R_1 xz]$, then by the identities just mentioned, we obtain the right condition, $x[\lambda z\, S_1 zy] \,\&\, y[\lambda z\, S_1 xz]$. And analogous reasoning establishes the right-to-left direction.

So it remains only to show $[\lambda z\, R_1 zy] = [\lambda z\, S_1 zy]$ and $[\lambda z\, R_1 xz] = [\lambda z\, S_1 xz]$. Note that since we have assumed that there is exactly one necessarily unexemplified property, it suffices to show (exercise) that the properties $[\lambda z\, R_1 zy]$,

[114] Consider $R_!$. It will be seen later that $[\lambda uv\; Fua \,\&\, Fvb]\downarrow$, where F, u, and v are variables, is axiomatic, by (39.2). Since F occurs free, the universal closure of this axiom is also axiomatic, namely, $\forall F([\lambda uv\; Fua \,\&\, Fvb]\downarrow)$. So since $=_E\downarrow$, we can instantiate $=_E$ into the universal claim and apply infix notation to obtain $[\lambda uv\; u=_E a \,\&\, v=_E a]\downarrow$. And similar reasoning applies to S_1.

$[\lambda z\, S_1 zy]$, $[\lambda z\, R_1 xz]$, and $[\lambda z\, S_1 xz]$ are all necessarily unexemplified. Without loss of generality, then, we show only that the first, namely $[\lambda z\, R_1 zy]$, is necessarily unexemplified, since analogous reasoning yields the conclusion for the other three.

Now it will be provable that $[\lambda z\, R_1 zy]$ exists given that R_1 exists. But by definition of R_1, this property just is $[\lambda z\, [\lambda uv\, u =_E a \,\&\, v =_E a]zy]$. Consequently, where w is any individual, strengthened β-Conversion (181) implies the following biconditional chain:

$$\begin{aligned} [\lambda z\, [\lambda uv\, u =_E a \,\&\, v =_E a]zy]w &\equiv [\lambda uv\, u =_E a \,\&\, v =_E a]wy \\ &\equiv w =_E a \,\&\, y =_E a \end{aligned}$$

But, by hypothesis, y is abstract. Thus, we know, by reasoning from the definition of $=_E$ (230), that $\neg(y =_E a)$. Hence, $\neg(w =_E a \,\&\, y =_E a)$. So given our biconditional chain, $\neg([\lambda z\, [\lambda uv\, u =_E a \,\&\, v =_E a]zy]w)$, i.e., $\neg[\lambda z\, R_1 zy]w$. This holds for *any* object w, i.e., $\forall w \neg[\lambda z\, R_1 zy]w$, i.e., $\neg\exists w[\lambda z\, R_1 zy]w$. Since we reasoned our way to this conclusion without appealing to any contingent assumptions, $\Box\neg\exists w[\lambda z\, R_1 zy]w$, i.e., $[\lambda z\, R_1 zy]$ is necessarily unexemplified.

Since analogous reasoning allows us to establish that the other properties $[\lambda z\, R_1 xz]$, $[\lambda z\, S_1 zy]$, and $[\lambda z\, S_1 xz]$ are necessarily unexemplified for abstract x and y, it follows from the hypothesis that there is at most one necessarily unexemplified property that $[\lambda z\, R_1 zy] = [\lambda z\, S_1 zy]$ and $[\lambda z\, R_1 xz] = [\lambda z\, S_1 xz]$. And this was all that remained to be shown to complete our proof that, under the definition of relation identity given by (ϑ), the hypothesis that there is a unique necessarily unexemplified property implies that there is at most one ordinary object.

One could argue that such a result is without consequence, since pretheoretically, it seems reasonable to assert that there are distinct, necessarily unexemplified properties, such as *being a barber who shaves all and only those who don't shave themselves* and *being a brown and colorless dog*. If we assert the existence of distinct such properties, as our system allows us to do, we avoid the result that there is at most one ordinary object. But it is somewhat curious that what seems like a natural definition of relation identity should force us to assert the existence of distinct such properties. Further investigation is certainly called for. Fortunately, we need not positively assert the existence of distinct such properties if we proceed by eschewing (ϑ) in favor of the more traditional definition (23.3).

Exercise: Give a reason for thinking that the official definition of relation identity (23.3) does not imply, when F and G exist, that if there is exactly one necessarily unexemplified property, then there is at most one ordinary object. As an outline of the reason, let F and G be any relations known to exist, so that (23.3) reduces to:

$(\omega)\ \ F = G \equiv \forall y([\lambda z\, Fzy] = [\lambda z\, Gzy]\ \&\ [\lambda z\, Fyz] = [\lambda z\, Gyz])$

Then one can show that this is equivalent to:[115]

$$F = G \equiv \forall x \forall y((x[\lambda z\, Fzy] \equiv x[\lambda z\, Gzy])\ \&\ (x[\lambda z\, Fyz] \equiv x[\lambda z\, Gyz]))$$

At this point, note that the above does lead to a claim analogous to the **Fact** we established for (ϑ). The **Fact** showed that (ϑ) yields identity conditions for relations that are independent of the properties that result when the relations projected to *ordinary* objects. But the identity conditions for relations given by (23.3) don't allow us to ignore such projections. It is not a fact that:

$$\forall x \forall y(\neg(A!x\ \&\ A!y) \to ((x[\lambda z\, Fzy] \equiv x[\lambda z\, Gzy])\ \&\ (x[\lambda z\, Fyz] \equiv x[\lambda z\, Gyz])))$$

An argument analogous to the one used to establish the **Fact** fails to establish the above claim.

[115] Here's why. By definition of property identity (23.2), (ω) is equivalent to:

$$F = G \equiv \forall y(\Box \forall x(x[\lambda z\, Fzy] \equiv x[\lambda z\, Gzy])\ \&\ \Box \forall x(x[\lambda z\, Fyz] \equiv x[\lambda z\, Gyz]))$$

Since $\Box\varphi\ \&\ \Box\psi$ is equivalent to $\Box(\varphi\ \&\ \psi)$, the above is equivalent to:

$$F = G \equiv \forall y \Box(\forall x(x[\lambda z\, Fzy] \equiv x[\lambda z\, Gzy])\ \&\ \forall x(x[\lambda z\, Fyz] \equiv x[\lambda z\, Gyz]))$$

By the Barcan Formula, this is equivalent to:

$$F = G \equiv \Box \forall y(\forall x(x[\lambda z\, Fzy] \equiv x[\lambda z\, Gzy])\ \&\ \forall x(x[\lambda z\, Fyz] \equiv x[\lambda z\, Gyz]))$$

And for reasons similar to those described in footnote 112, this is equivalent to:

$$F = G \equiv \forall y(\forall x(x[\lambda z\, Fzy] \equiv x[\lambda z\, Gzy])\ \&\ \forall x(x[\lambda z\, Fyz] \equiv x[\lambda z\, Gyz]))$$

Since $\forall x\varphi\ \&\ \forall x\psi$ is equivalent to $\forall x(\varphi\ \&\ \psi)$, the above is equivalent to:

$$F = G \equiv \forall y \forall x((x[\lambda z\, Fzy] \equiv x[\lambda z\, Gzy])\ \&\ (x[\lambda z\, Fyz] \equiv x[\lambda z\, Gyz]))$$

Finally, since universal quantifiers commute, the above is equivalent to:

$$F = G \equiv \forall x \forall y((x[\lambda z\, Fzy] \equiv x[\lambda z\, Gzy])\ \&\ (x[\lambda z\, Fyz] \equiv x[\lambda z\, Gyz])) \qquad \bowtie$$

Chapter 8

Axioms

Now that we have a precisely-specified philosophical language that allows to express claims using primitive and defined notions, we next assert the fundamental axioms of our theory in terms of these notions. We may group these axioms as follows:

- Axioms for negations and conditionals.
- Axioms for universal quantification and logical existence.
- Axioms for the substitution of identicals.
- Axioms for actuality.
- Axioms for necessity.
- Axioms for necessity and actuality.
- Axioms for definite descriptions.
- Axioms for relations.
- Axioms for encoding.

The statement of certain axiom groups require preliminary definitions.

In what follows, when we assert the closures of a *schema* as axioms, we mean that the closures of every instance of the schema are axioms. Moreover, it is important to recognize that when axioms are expressed not as schemata that use metavariables, but as statements involving object-language variables, these variables are *not* functioning as metavariables. Convention (17.2) applies only to definitions, *not* to axioms and theorems.

8.1 Axioms for Negations and Conditionals

(**38**) **Axioms**: Negations and Conditionals. To ensure that negation and conditionalization behave classically, we take the closures of the following schemata as axioms:

(.1) $\varphi \to (\psi \to \varphi)$

(.2) $(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi))$

(.3) $(\neg\varphi \to \neg\psi) \to ((\neg\varphi \to \psi) \to \varphi)$

By using the above, instead of defining *tautology* and asserting that all tautologies are axioms, we avoid the use of semantic notions in the statement of the system.

8.2 Axioms for Universal Quantification and Logical Existence

The axioms below assume familiarity with the syntactic notions *primitive constant* (1), *variable* (1), *matrix* (3.8), *subformula* (6), *primary term* (7.8), *free/bound* variables (8), *encoding position* (9.1), *core* λ-expressions (9.2), and *substitutable for* (15). These definitions were developed in the previous chapter, along with the definitions upon which they depend. For example, it should be clear what is meant when (39.2) references *core* λ-expressions, namely, that no variable bound by the λ occurs in encoding position in the matrix.

(**39**) **Axioms:** Quantification and Logical Existence. Where α is any variable, τ is any term, Π^n is any n-ary relation term, and $\kappa_1,\ldots,\kappa_n$ are any individual terms, we assert the closures of the following as axioms:

(.1) $\forall\alpha\varphi \to (\tau{\downarrow} \to \varphi^{\tau}_{\alpha})$, provided τ is substitutable for α in φ

(.2) $\tau{\downarrow}$, whenever τ is either a primitive constant (i.e., one not introduced by a definition), a variable, or a core λ-expression.

(.3) $\forall\alpha(\varphi \to \psi) \to (\forall\alpha\varphi \to \forall\alpha\psi)$

(.4) $\varphi \to \forall\alpha\varphi$, provided α doesn't occur free in φ

(.5) (a) $\Pi^n\kappa_1\ldots\kappa_n \to (\Pi^n{\downarrow} \;\&\; \kappa_1{\downarrow} \;\&\; \ldots \;\&\; \kappa_n{\downarrow})$,
where $\Pi^n\kappa_1\ldots\kappa_n$ is any exemplification formula ($n \geq 0$)
(b) $\kappa_1\ldots\kappa_n\Pi^n \to (\Pi^n{\downarrow} \;\&\; \kappa_1{\downarrow} \;\&\; \ldots \;\&\; \kappa_n{\downarrow})$,
where $\kappa_1\ldots\kappa_n\Pi^n$ is any encoding formula ($n \geq 1$)

As we shall see, these axioms (i) yield classical quantification theory with respect to primitive constants, variables, and core λ-expressions, and (ii) yield a negative free logic of quantification for reasoning with respect to all other terms. (.1) asserts that significant terms can be instantiated into universal claims;[116] (.2) asserts the significance of primitive constants, variables, and core λ-expressions; (.3) and (.4) are part of classical quantification theory; and (.5.a) and (.5.b) assert not only that the primary terms of true exemplification and encoding formulas are significant, but also that true formulas are significant (this is the 0-ary case of (.5.a), since formulas are 0-ary relation terms).

The reader may find the observations in the next Remark useful for understanding (.2) in more depth. See also Remark (55) at the end of this chapter.

(**40**) **Remark:** Digression on the Formulation of (39.2). (39.2) asserts the significance of primitive constants, variables, and core λ-expressions. Later, we'll derive that every formula, i.e., every 0-ary relation term, is significant (104.2).

We've already discussed, in the third observation in Remark (25), why (39.2) asserts that *primitive* constants, as opposed to constants generally, are significant. (39.2) does *not* assert that constants introduced by definition are significant. The loss is not an egregious one. If we have defined a new constant τ by means of a closed complex term that is *provably significant*, then the inferential role of definitions-by-identity, sketched earlier in (17) and articulated precisely in (73) and (120), ensures that we can derive $\tau\downarrow$; a definition-by-= implies that if the definiens of a new constant is significant, then the identity stipulated in the definition holds. In our system, true identity claims will imply the existence of the entities identified – see the theorems in (107) – and so there will be a means of deriving the significance of new constants introduced by significant, closed terms in a definition-by-identity.

There should be no question as to why (39.2) stipulates that variables are significant; in a logic in which open formulas are assertible, free variables are assumed to have an arbitrary value. So, (39.2) tells us that we may assert $\alpha\downarrow$ for any variable α. We can always instantiate the variable α into the universal claim $\forall\beta\varphi$ provided α is substitutable for β in φ.

It should be clear why (39.2) doesn't assert $\tau\downarrow$ when τ is a definite description. Some definite descriptions may contingently fail to be significant, while others, such as $\imath x(Fx \,\&\, \neg Fx)$, provably fail to have a denotation (for every property F). Of course, some definite descriptions will provably be significant. We'll see that $\imath x(x=y)\downarrow$ in (177) and that $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))\downarrow$, when x doesn't occur free in φ (252).

[116] It should be kept in mind that the formula $\tau\downarrow$ used in (.1) and (.2) is defined differently for individual terms and relation terms. So what the instances of $\tau\downarrow$ assert when they are expanded by definition depends on whether τ is an individual term or a relation term, and if the latter, the arity of the relation.

Finally, we discuss what (39.2) asserts in connection with λ-expressions. First, it should be clear that the following are all axiomatic according to (39.2):

- $[\lambda x\, \iota y(Qyx)P]\downarrow$
- $[\lambda x\, \neg\iota y(Qyx)P]\downarrow$
- $[\lambda x\, y[\lambda z\, Rxz]]\downarrow$
- $[\lambda x\, \Box\forall z(z[\lambda y\, Fyx] \equiv z[\lambda y\, Gyx])]$
- $[\lambda x\, \iota y(Gx \equiv \neg yG)P]\downarrow$
- $[\lambda x\, \iota y(Gx \equiv yG)P]\downarrow$
- $[\lambda x\, [\lambda y\, \exists H(yH \,\&\, \neg Hy)]x]\downarrow$

In these claims, the complex term is a core λ-expression (9.2): the λ does not bind a variable that occurs in encoding position in the matrix (9.1).

Second, to see why (39.2) only asserts the significance of core λ-expressions, consider the fact that in classical second-order (exemplification) logic, the comprehension condition for relations is unrestricted; for every formula φ, the λ-expression $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ is significant. But in second-order logic with encoding formulas, the paradoxes of encoding preclude the simultaneous assertion of an unrestricted comprehension principle for relations and an unrestricted comprehension principle for abstract objects. We therefore restrict comprehension for relations while maintaining unrestricted comprehension for abstract objects. The intuitive idea is to avoid asserting, in the first instance, the existence of relations that would make an encoding condition an integral rather than an incidental part of the exemplification conditions of the relation. (39.2) captures this intuition mentioned in the previous chapter: if the λ in $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ were to bind a variable that occurs in encoding position in φ, then it would effectively make an encoding condition an integral rather than an incidental part of the exemplification conditions of the relation denoted.

Now, with this understanding, there are some interesting facts about the variety of core λ-expressions that are covered by (39.2). Note that the axiom doesn't require that core λ-expressions contain only primitive terms; core λ-expressions may be significant even if they contain occurrences of defined terms. Indeed, in some cases, they may even contain occurrences of *empty* defined terms (i.e., terms introduced in a definition-by-= in which the definiens is an empty term). We can see this by considering the examples discussed in the penultimate paragraph of (17). Axiom (39.2) asserts that the core expression $[\lambda x\, Rx\iota z(Pz \& \neg Pz)]$ is significant, despite having an empty description as a subterm. The open formula $Rx\iota z(Pz \,\&\, \neg Pz)$ is necessarily false for every x, and so the λ-expression will denote a property that is necessarily unexemplified. So if we define $a =_{df} \imath z(Pz \,\&\, \neg Pz)$, then (39.2) asserts that $[\lambda x\, Rxa]\downarrow$.

The second example in the penultimate paragraph of (17) is also relevant; in this case, we have a core λ-expression that contains an empty λ-expression. Start with the λ-expression that plays a role in the Clark/Boolos paradox, namely, $[\lambda x\, \exists G(xG \,\&\, \neg Gx)]$; this is provably empty on pain of contradiction, as established in theorem (192.1). Now consider the following expression: $[\lambda y\, [\lambda x \exists G(xG \& \neg Gx)]y]$. This is a core λ-expression (9.2), since y doesn't occur in encoding position in the matrix (9.1). So (39.2) asserts $[\lambda y [\lambda x \exists G(xG \& \neg Gx)]y]\downarrow$. And if one were to define $Q =_{df} [\lambda x\, \exists G(xG \,\&\, \neg Gx)]$, then (39.2) would assert $[\lambda y\, Qy]\downarrow$. In Remark (155) of Chapter 9, we'll see a number of examples of descriptions and λ-expressions that are provably significant despite the fact that they contain empty subterms.

Although (39.2) asserts that only core λ-expressions are significant, we may still prove that some non-core λ-expressions are significant. For example, a non-core λ-expression will be provably significant if it has a matrix that is necessarily and universally equivalent to the matrix of a λ-expression that is known to be significant. This will be systematized as axiom (49). And the Kirchner Theorem (271.2) will also state conditions under which non-core λ-expressions are provably significant. Of course, some non-core λ-expressions, such as those leading to the Clark/Boolos, McMichael/Boolos, and Kirchner paradoxes, will be provably empty.

8.3 Axioms for the Substitution of Identicals

(**41**) **Axioms:** The Substitution of Identicals. The identity symbol '=' is not a primitive expression of our object language. Instead, identity was defined in items (23.1), (23.2), (23.3), and (23.4) for individuals and n-ary relations ($n \geq 0$). The classical law of the reflexivity of identity, i.e., $\alpha = \alpha$, where α is any variable, will be derived as a theorem — see item (117.1) in Chapter 9. By contrast, we take the classical law of the substitution of identicals as an axiom. The closures of the following schema are therefore asserted as axioms of our system:

> $\alpha = \beta \rightarrow (\varphi \rightarrow \varphi')$, whenever β is substitutable for α in φ, and φ' is the result of replacing zero or more free occurrences of α in φ with occurrences of β.

This is an unrestricted principle of substitution of identicals: (a) if x and y are identical individuals, then anything we can prove about x holds of y, and (b) if F^n and G^n are identical n-ary relations, then anything we can prove about F^n holds of G^n.

8.4 Axioms for Actuality

(42) **Metadefinition:** Modally Fragile Axioms. In (43), we shall introduce a special group of axioms that have a profound affect on modal reasoning within our system. We call these *modally fragile* axioms. We define:

> φ is *modally fragile* just in case φ is (asserted as) an axiom, but its modal closures are not.

There are various material or formal reasons why we may want to assert φ, but not its modal closures, as axioms. Materially, we may know or believe that φ is true but its necessitation is not, such as when φ is a contingent truth knowable only *a posteriori*. Formally, it may be that φ is semantically valid but $\Box\varphi$ is not, such as when φ is a contingent truth knowable *a priori*. So any contingent truth added as an axiom would be classified as modally fragile. We'll see later on that the notion of modal fragility also applies to a claim φ that is discoverable only *a posteriori* but whose necessity can be subsequently established once φ is asserted as an axiom or assumption. In these cases, we refrain from asserting the necessitation of φ as an axiom, notwithstanding its subsequent derivation as a theorem. (For an example of these latter cases, see (137).) In all these cases, we shall say that the axioms in question are modally fragile, in the above sense.

Consequently, in what follows, we use the label '$\star$**Axiom**' to signpost that an axiom is modally fragile and subsequently place a $\star$ adjacent to its item number whenever we cite it. The $\star$ adjacent to the *axiom* number thereby signals that the modal closures of that principle are not asserted as axioms. (As we shall see, a $\star$ adjacent to a *theorem* number has a slightly different, but related, meaning.) When we reason from axioms that are modally fragile, we have to be especially careful when applying the Rule of Necessitation. Our proof theory will be formulated so that *if* the proof of a theorem φ contains an inferential step that is justified by a modally fragile axiom, then we may not use the Rule of Necessitation to infer that $\Box\varphi$ is also a theorem (though we may be able to infer $\Box\varphi$ by other means).

(43) $\star$**Axioms:** Modally Fragile Axioms of Actuality. Where φ is any formula, we assert *only the universal closures* of the following as modally fragile axioms of the system:

$$\mathcal{A}\varphi \rightarrow \varphi$$

The necessitation of the above schema, i.e., $\Box(\mathcal{A}\varphi \rightarrow \varphi)$, is not semantically valid, and so we do not assert the modal closures of the schema's instances as axioms.[117] That's why this is designated as a $\star$-axiom. Instances of this axiom

[117] Intuitively, just consider models of the theory in which there are two possible worlds, $\boldsymbol{w}_\alpha$ and

schema are logical truths that aren't necessary as well as contingent truths that are knowable *a priori* (Zalta 1988b, 73).

We need not assert the biconditional $\mathcal{A}\varphi \equiv \varphi$ as a modally fragile axiom since the right-to-left direction, $\varphi \to \mathcal{A}\varphi$, will be derivable by non-modally strict means as a ⋆-theorem (130.1)⋆.[118] Moreover, it is important to note that the *actualization* of the above axiom, i.e., $\mathcal{A}(\mathcal{A}\varphi \to \varphi)$ is derivable as a modally strict theorem (133.1); as is $\mathcal{A}(\varphi \to \mathcal{A}\varphi)$ (133.2) and $\mathcal{A}(\mathcal{A}\varphi \equiv \varphi)$ (133.4). Finally, where a $\Box$-free closure is defined as in (11), we'll see that all the $\Box$-free closures of $\mathcal{A}\varphi \to \varphi$ are derivable by modally strict means (134).

(44) **Axioms:** Necessary Axioms of Actuality. By constrast we take the all of the closures of the following axiom schemata to be axioms of the system:

(.1) $\mathcal{A}\neg\varphi \equiv \neg\mathcal{A}\varphi$

(.2) $\mathcal{A}(\varphi \to \psi) \equiv (\mathcal{A}\varphi \to \mathcal{A}\psi)$

(.3) $\mathcal{A}\forall\alpha\varphi \equiv \forall\alpha\mathcal{A}\varphi$

(.4) $\mathcal{A}\varphi \equiv \mathcal{A}\mathcal{A}\varphi$

We leave the axioms governing the interaction between the actuality operator and the necessity operator for Section 8.6.[119]

8.5 Axioms for Necessity

(45) **Axioms:** Necessity. We take the closures of the following principles as axioms:

(.1) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$ (K)

(.2) $\Box\varphi \to \varphi$ (T)

(.3) $\Diamond\varphi \to \Box\Diamond\varphi$ (5)

Furthermore, if we read the formula $\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$ as asserting the existence of a concrete-but-not-actually-concrete object, then the following axiom asserts that it is *possible* that such objects exist:

$\boldsymbol{w}_1$. In any such model where φ is true at the actual world $\boldsymbol{w}_\alpha$ but false at $\boldsymbol{w}_1$, then $\Box(\mathcal{A}\varphi \to \varphi)$ is false at $\boldsymbol{w}_\alpha$ because the conditional $\mathcal{A}\varphi \to \varphi$ is false at $\boldsymbol{w}_1$ (at $\boldsymbol{w}_1$, the antecedent is true and the consequent is false).

[118] Previous versions of this monograph took the biconditional $\mathcal{A}\varphi \equiv \varphi$ as axiomatic. But I had not properly done my homework; I thank Daniel Kirchner for pointing out that the left-to-right condition of the biconditional suffices as an axiom.

[119] In thinking about the logic of actuality, I've benefited from reading Hazen 1978, 1990, and Hazen, Rin, & Wehmeier 2013.

(.4) $\Diamond\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$

(.1) – (.3) are well known. Axiom (.4) may not be familiar, but it will play a significant role in what follows. We'll discuss it in more detail on several occasions. For now, it suffices to say that its philosophical justification is that it captures the idea that the world might have contained something distinct from every actually concrete thing. Interested readers may wish to examine Remark (56), in which reasons are given for preferring (.4) to alternative axioms that might suggest themselves.

Whereas many systematizations of quantified S5 modal logic use a primitive Rule of Necessitation (RN), we derive RN in item (68) (Chapter 9). RN is derived in a form that guarantees that if there is a proof of a formula φ that doesn't depend on a modally-fragile axiom, then there is a proof of $\Box\varphi$. Thus, we can't apply RN to any formula derived from the logic of actuality axiom (43)★. Once RN is derived, we shall be able to derive the Barcan Formula and Converse Barcan Formula; this occurs in item (168).

8.6 Axioms for Necessity and Actuality

(**46**) **Axioms:** Necessity and Actuality. We take the closures of the following principles as axioms:

(.1) $\mathcal{A}\varphi \rightarrow \Box\mathcal{A}\varphi$

(.2) $\Box\varphi \equiv \mathcal{A}\Box\varphi$

(.1) asserts that if it is actually the case that φ, then necessarily it is actually the case that φ; (.2) asserts that φ is necessary if and only if it is actually the case that φ is necessary.

8.7 Axioms for Definite Descriptions

(**47**) **Axioms:** Descriptions. We take the closures of the following axiom schema as axioms:

> $x = \imath x\varphi \;\equiv\; \forall z(\mathcal{A}\varphi^z_x \equiv z = x)$, provided z is substitutable for x in φ and doesn't occur free in φ

We may read this as: x is identical to the individual that is (in fact) such that φ just in case all and only those individuals that are actually such that φ are identical to x. The notation $\mathcal{A}\varphi^z_x$ used in the axiom involves a harmless ambiguity; it should strictly be formulated as $(\mathcal{A}\varphi)^\tau_\nu$. But by the third bullet point in definition (14), we know that $(\mathcal{A}\varphi)^\tau_\nu = \mathcal{A}(\varphi^\tau_\nu)$.

8.8 Axioms for Relations

(**48**) **Axioms:** Complex n-ary Relation Terms. The axioms traditionally labeled α-Conversion and β-Conversion govern the *significant* λ-expressions of the form $[\lambda x_1 \dots x_n\, \varphi]$, while η-Conversion governs only *elementary* λ-expressions of the form $[\lambda x_1 \dots x_n\, F^n x_1 \dots x_n]$:

(.1) α-Conversion. Where $\nu_1, \dots, \nu_n$ are any distinct individual variables, the closures of the following are axioms ($n \geq 0$):

$$[\lambda \nu_1 \dots \nu_n\, \varphi]\downarrow \to [\lambda \nu_1 \dots \nu_n\, \varphi] = [\lambda \nu_1 \dots \nu_n\, \varphi]',$$
where $[\lambda \nu_1 \dots \nu_n\, \varphi]'$ is any alphabetic variant of $[\lambda \nu_1 \dots \nu_n\, \varphi]$

(.2) β-Conversion. The closures of the following are axioms ($n \geq 1$):

$$[\lambda x_1 \dots x_n\, \varphi]\downarrow \to ([\lambda x_1 \dots x_n\, \varphi]x_1 \dots x_n \equiv \varphi)$$

(.3) η-Conversion: Where $n \geq 0$, the closures of the following are axioms:

$$[\lambda x_1 \dots x_n\, F^n x_1 \dots x_n] = F^n$$

α-Conversion guarantees that significant, alphabetically-variant n-ary λ-expressions ($n \geq 0$) denote the same relation. We shall see that a stronger, *unconditional*, version of α-Conversion can be derived for 0-ary λ-expressions; see (111.3) and (111.4).

Note that α-Conversion is formulated entirely with metavariables, while β-Conversion is formulated with a mixture of object-language variables and metavariables, and η-Conversion is formulated entirely in terms of object-language variables. Thus, α-Conversion governs any λ-expressions having the right form.

By contrast, β-Conversion is formulated with respect to (a) λ-expressions in which the λ binds the distinct object-language variables $x_1, \dots, x_n$ and (b) exemplification formulas involving the individual variables $x_1, \dots, x_n$. We've formulated β-Conversion with specific object-language variables because we can later *prove* a strengthened form of β-Conversion (181). It asserts that $[\lambda \mu_1 \dots \mu_n\, \varphi]\downarrow \to ([\lambda \mu_1 \dots \mu_n\, \varphi]\nu_1 \dots \nu_n \equiv \varphi^{\nu_1, \dots, \nu_n}_{\mu_1, \dots, \mu_n})$, for any distinct individual variables $\mu_1, \dots, \mu_n$ and for any individual variables $\nu_1, \dots, \nu_n$, provided $\nu_1, \dots, \nu_n$ are substitutable, respectively, for $\mu_1, \dots, \mu_n$ in φ.[120]

[120] It is worth observing an interesting fact about β-Conversion and λ-expressions that are built from empty λ-expressions. The λ-expression $[\lambda z\, \exists G(zG \,\&\, \neg Gz)]$ is empty (it leads to the Clark-Boolos paradox) but the following λ-expression provably exists:

$$[\lambda x\, [\lambda z\, \exists G(zG \,\&\, \neg Gz)]x]$$

By the discussion at the end of (39), we know that $[\lambda x\, [\lambda z \exists G(zG \,\&\, \neg Gz)]x]\downarrow$ is an instance of axiom (39.2). This establishes the antecedent of the following instance of β-Conversion:

(ϑ) $[\lambda x\, [\lambda z\, \exists G(zG \,\&\, \neg Gz)]x]\downarrow \to ([\lambda x\, [\lambda z\, \exists G(zG \,\&\, \neg Gz)]x]x \equiv [\lambda z\, \exists G(zG \,\&\, \neg Gz)]x)$

β-Conversion (.2) guarantees that significant n-ary λ-expressions behave classically: the left-to-right direction of the consequent is sometimes referred to as λ-Conversion, while the right-to-left direction is sometimes referred to as λ-Abstraction. We don't need to assert β-Conversion for 0-ary λ-expressions of the form $[\lambda\,\varphi]$ because it follows from theorem (111.2), where we derive the *unconditional* claim $[\lambda\,\varphi] \equiv \varphi$ for any formula φ, *without* appealing to the 0-ary case of (.2). Since $[\lambda\,\varphi] \equiv \varphi$ is derivable without appeal to the 0-ary case of (.2), the 0-ary case of (.2) becomes a theorem, by the axiom of propositional logic $\chi \to (\psi \to \chi)$.

Finally, η-Conversion is not schematic at all. It governs the λ-expressions of the form $[\lambda x_1 \dots x_n\, F^n x_1 \dots x_n]$, for any $n \geq 0$. These expressions have, as a matrix, an n-ary exemplification formula $F^n x_1 \dots x_n$ (with nothing but free variables), and where the x_i ($0 \leq i \leq n$) are all bound by the λ. Since the universal closures of $[\lambda x_1 \dots x_n\, F^n x_1 \dots x_n] = F^n$ with respect to the sole free variable F are axiomatic, the negative free logic of quantification will guarantee that only significant λ-expressions can be instantiated into such universal closures to produce instances. For example, the following is a universal closure of the unary case of η-Conversion and so an axiom:

$$\forall F([\lambda x\, Fx] = F)$$

Since the λ-expression $[\lambda z\, \exists G(zG \,\&\, \neg Gz)]$ isn't significant, we won't be able to instantiate it into the above universal claim to derive:

$$[\lambda x\, [\lambda z\, \exists G(zG \,\&\, \neg Gz)]x] = [\lambda z\, \exists G(zG \,\&\, \neg Gz)]$$

This can't be a theorem, since the term on the left of the identity symbol is significant (for the reasons mentioned in footnote 120) while the term on the right is not (on pain of the Clark/Boolos paradox).

However, we later prove theorems that extend η-Conversion in two ways: (a) in (111.1), we derive, for $n = 0$, an unconditional equation of the form $[\lambda\,\varphi] = \varphi$, and (b) in (186.2), we derive $\Pi^n\!\downarrow \to [\lambda \nu_1 \dots \nu_n\, \Pi^n \nu_1 \dots \nu_n] = \Pi^n$ for $n \geq 0$, where $[\lambda \nu_1 \dots \nu_n\, \Pi^n \nu_1 \dots \nu_n]$ is any elementary λ-expression, Π^n is any n-ary relation term, and $\nu_1, \dots, \nu_n$ are any distinct object variables that have no free occurrences in Π^n.

(**49**) **Axioms:** Coexistence of Relations. We now assert that if the relation *being* $x_1 \dots x_n$ *such that* φ exists and necessarily, for all $x_1, \dots, x_n$, φ is materially equivalent to ψ, then the relation *being* $x_1 \dots x_n$ *such that* ψ exists ($n \geq 1$):

Hence it follows that:

(ξ) $[\lambda x\, [\lambda z\, \exists G(zG \,\&\, \neg Gz)]x]x \equiv [\lambda z\, \exists G(zG \,\&\, \neg Gz)]x$

But since the Clark/Boolos expression $[\lambda z\, \exists G(zG \,\&\, \neg Gz)]$ can't be significant on pain of contradiction, it will be provable (192.1) that $\neg[\lambda z\, \exists G(zG \,\&\, \neg Gz)]\!\downarrow$. Hence, it follows by axiom (39.5.a) that $\neg[\lambda z\, \exists G(zG \,\&\, \neg Gz)]x$. So from this and ($\xi$), it follows that $\neg[\lambda x\, [\lambda z\, \exists G(zG \,\&\, \neg Gz)]x]x$. No x exemplifies the property signified by $[\lambda x\, [\lambda z\, \exists G(zG \,\&\, \neg Gz)]x]$.

$$([\lambda x_1 \ldots x_n\, \varphi]\!\downarrow \,\&\, \Box\forall x_1 \ldots \forall x_n(\varphi \equiv \psi)) \rightarrow [\lambda x_1 \ldots x_n\, \psi]\!\downarrow \qquad (n \geq 1)$$

Intuitively, this tells us that if a λ-expression is significant, then any λ-expression with a necessarily and universally equivalent matrix is also significant. We don't need to assert the case when $n=0$, since $[\lambda\, \psi]\!\downarrow$ is an instance of axiom (39.2) and the 0-ary case, $[\lambda\, \varphi]\!\downarrow \,\&\, \Box(\varphi \equiv \psi) \rightarrow [\lambda\, \psi]\!\downarrow$, follows by the truth of the consequent.

We also note, for future reference, that this axiom is *not* a trivial instance or consequence of the Rules of Substitution that we derive in item (160).[121] Some readers may be interested in the discussion Remark (57), where we explain why the necessity operator and universal quantifiers are needed in the second conjunct of the antecedent of the above axiom.

It is worth mentioning again that this axiom allows us to prove the significance of some λ-expression *excluded* by (39.2). Consider a λ-expression Π in which the initial λ binds a variable that occurs in encoding position in the matrix. If one can show that the matrix of Π is necessarily and universally equivalent to the matrix of a λ-expression whose significance is known, then we can infer that Π is significant. To see an example, examine the definiens used in (230), namely, $[\lambda xy\, O!x \,\&\, O!y \,\&\, x = y]$. In this expression, the λ binds variables that occur in encoding position in the matrix – by Convention (17.3), x and y occur in encoding position in $x = y$. However, theorem (229) establishes that $[\lambda xy\, O!x \,\&\, O!y \,\&\, x = y]$ exists, and the proof in the Appendix does this by showing that the matrix is necessarily and universally equivalent to the matrix of a significant λ-expression, namely, $[\lambda xy\, O!x \,\&\, O!y \,\&\, x =_E y]$.

8.9 Axioms for Encoding

(**50**) **Axiom:** n-ary Encoding Reduction. When $n \geq 2$, we assert that individuals $x_1, \ldots, x_n$ encode relation F^n if and only if: x_1 encodes $[\lambda y\, F^n y x_2 \ldots x_n]$ and x_2 encodes $[\lambda y\, F^n x_1 y x_3 \ldots x_n]$ and … and $x_n[\lambda y\, F^n x_1 \ldots x_{n-1} y]$. Formally, we assert the closures of:

$$\begin{aligned} &x_1 \ldots x_n F^n \equiv \\ &\quad x_1[\lambda y\, F^n y x_2 \ldots x_n] \,\&\, x_2[\lambda y\, F^n x_1 y x_3 \ldots x_n] \,\&\, \ldots \,\&\, x_n[\lambda y\, F^n x_1 \ldots x_{n-1} y] \end{aligned}$$

[121] In the next chapter, we introduce Rules of Substitution and the notion of *modal strictness*. The Rules of Substitution (160) allow one to substitute, within a derivation, a formula ψ for a formula φ only when (a) φ occurs as a subformula of some formula, say χ, and (b) there is either a proof of $\Box(\varphi \equiv \psi)$ or a *modally strict* proof (60.2) of $\varphi \equiv \psi$. The above axiom isn't implied by these rules, since by definition (6), φ is not a subformula of the term $[\lambda x_1 \ldots x_n\, \varphi]$ and so is not a subformula of the claim $[\lambda x_1 \ldots x_n\, \varphi]\!\downarrow$. Moreover, unlike the Rules of Substitution, instances of the above axiom don't require a *proof* of $\Box(\varphi \equiv \psi)$ or a modally strict proof of $\varphi \equiv \psi$; if $\Box\forall x_1 \ldots \forall x_n(\varphi \equiv \psi)$ were only an assumption, for example, we could use that fact when applying the above axiom. See the discussion in (71). For further discussion of some of these issues, see the examples of illegitimate uses of the Rules of Substitution in (161).

where $n \geq 2$.

It is natural to ask here: why not just formulate object theory using just unary encoding predications of the form xF and then use the above to *define* n-ary encoding predications as a conjunction of unary encoding predications? The answer has to do with one of the motivations underlying object, namely, to give sentences like 'Holmes is a friend of Watson' and '1 < 2' a true reading on which they are binary predications (in this case, the true reading is a binary encoding predication) and not a conjunction, just as 'Holmes is a detective' and '2 is prime' have a true reading on which they are unary encoding predications.

(**51**) **Axiom:** Rigidity of Encoding. If an object x encodes a property F, it does so necessarily. That is, the closures of the following are axioms of the system:

$$xF \to \Box xF$$

In other words, encoded properties are rigidly encoded. From this axiom, we will be able to prove both that $\Diamond xF \to \Box xF$ and $\Diamond xF \equiv \Box xF$. Moreover, given axiom (50), we shall be able to prove both that $x_1 \ldots x_n F^n \to \Box x_1 \ldots x_n F^n$ and $x_1 \ldots x_n F^n \equiv \Box x_1 \ldots x_n F^n$, for $n \geq 1$.

(**52**) **Axiom:** Ordinary Objects Fail to Encode Properties. The closures of the following are axioms:

$$O!x \to \neg \exists F\, xF$$

We prove in the next chapter that if x is ordinary, then *necessarily* x fails to encode any properties, i.e., that $O!x \to \Box\neg\exists F\, xF$.

(**53**) **Axioms:** Comprehension Principle for Abstract Objects ('Object Comprehension'). The closures of the following schema are axioms:

$$\exists x(A!x \,\&\, \forall F(xF \equiv \varphi)), \text{ provided } x \text{ doesn't occur free in } \varphi$$

When x doesn't occur free in φ, we may think of φ as presenting a condition on properties F, whether or not F is free in φ (the condition φ being a vacuous one when F doesn't occur free). So this axiom guarantees that for every condition φ on properties F expressible in the language, there exists an abstract object x that encodes just the properties F such that φ.

(**54**) **Remark:** The Restriction on Comprehension. In the formulation of the Comprehension Principle for Abstract Objects in (53), the formula φ used in comprehension may not contain free occurrences of x. This is a traditional constraint on comprehension schemata. Without such a restriction, a contradiction would be immediately derivable by using the formula '$\neg xF$' as φ, so as to produce the instance:

$$\exists x(A!x \,\&\, \forall F(xF \equiv \neg xF))$$

Any such object, say a, would be such that $\forall F(aF \equiv \neg aF)$, and a contradiction of the form $\varphi \equiv \neg\varphi$ would follow once we instantiate the quantifier $\forall F$ to any property. Instances such as the one displayed above are therefore ruled out by the restriction.

8.10 Summary of the Axioms

> *In order to investigate a subfield of a science, one bases it on the smallest possible number of principles, which are to be as simple, intuitive, and comprehensible as possible, and which one collects together and sets up as axioms.*
>
> — Hilbert 1922 (translated in Ewald 1996, 1119)

Hilbert's observation requires one minor amendment: one should base a science on the smallest number of principles *required to systematize the primitive and defined notions expressed in the language of the science*. With the help of the defined notions &, $\equiv$, $\exists$, $O!$, $A!$, $\downarrow$ and $=$, we have axiomatized the primitive notions $F^n x_1 \ldots x_n$, $x_1 \ldots x_n F^n$, $E!$, $\neg$, $\rightarrow$, $\forall$, $\Box$, $\mathcal{A}$, ι, and λ by asserting the closures of the following principles, with the exception of (43)★, of which only the universal closures are asserted.

Axioms for Negations and Conditionals:

- $\varphi \rightarrow (\psi \rightarrow \varphi)$ (38.1)
- $(\varphi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \chi))$ (38.2)
- $(\neg\varphi \rightarrow \neg\psi) \rightarrow ((\neg\varphi \rightarrow \psi) \rightarrow \varphi)$ (38.3)

Axioms for Universal Quantification and Logical Existence:

- $\forall\alpha\varphi \rightarrow (\tau\downarrow \rightarrow \varphi^\tau_\alpha)$, provided τ is substitutable for α in φ (39.1)
- $\tau\downarrow$, provided τ is primitive constant, a variable, or a core λ-expression (39.2)
- $\forall\alpha(\varphi \rightarrow \psi) \rightarrow (\forall\alpha\varphi \rightarrow \forall\alpha\psi)$ (39.3)
- $\varphi \rightarrow \forall\alpha\varphi$, provided α doesn't occur free in φ (39.4)
- $\Pi^n\kappa_1 \ldots \kappa_n \rightarrow (\Pi^n\downarrow \,\&\, \kappa_1\downarrow \,\&\, \ldots \,\&\, \kappa_n\downarrow) \quad (n \geq 0)$ (39.5.a)
  $\kappa_1 \ldots \kappa_n\Pi^n \rightarrow (\Pi^n\downarrow \,\&\, \kappa_1\downarrow \,\&\, \ldots \,\&\, \kappa_n\downarrow) \quad (n \geq 1)$ (39.5.b)

Axioms for the Substitution of Identicals:

- $\alpha = \beta \rightarrow (\varphi \rightarrow \varphi')$, whenever β is substitutable for α in φ, and φ' is the result of replacing zero or more free occurrences of α in φ with occurrences of β (41)

Axioms for Actuality:

- $\mathcal{A}\varphi \to \varphi$ (only universal closures) (43)★
- $\mathcal{A}\neg\varphi \equiv \neg\mathcal{A}\varphi$ (44.1)
- $\mathcal{A}(\varphi \to \psi) \equiv (\mathcal{A}\varphi \to \mathcal{A}\psi)$ (44.2)
- $\mathcal{A}\forall\alpha\varphi \equiv \forall\alpha\mathcal{A}\varphi$ (44.3)
- $\mathcal{A}\varphi \equiv \mathcal{A}\mathcal{A}\varphi$ (44.4)

Axioms for Necessity:

- $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$ (45.1)
- $\Box\varphi \to \varphi$ (45.2)
- $\Diamond\varphi \to \Box\Diamond\varphi$ (45.3)
- $\Diamond\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$ (45.4)

Axioms for Necessity and Actuality:

- $\mathcal{A}\varphi \to \Box\mathcal{A}\varphi$ (46.1)
- $\Box\varphi \equiv \mathcal{A}\Box\varphi$ (46.2)

Axioms for Definite Descriptions:

- $x = \imath x\varphi \;\equiv\; \forall z(\mathcal{A}\varphi^z_x \equiv z = x)$, provided z is substitutable for x in φ and doesn't occur free in φ (47)

Axioms for Relations:

- $[\lambda \nu_1 \ldots \nu_n\, \varphi]\!\downarrow\, \to [\lambda \nu_1 \ldots \nu_n\, \varphi] = [\lambda \nu_1 \ldots \nu_n\, \varphi]'$, where $n \geq 0$ and $[\lambda \nu_1 \ldots \nu_n\, \varphi]'$ is any alphabetic variant of $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ (48.1)
- $[\lambda x_1 \ldots x_n\, \varphi]\!\downarrow\, \to ([\lambda x_1 \ldots x_n\, \varphi]x_1 \ldots x_n \equiv \varphi)$ (48.2)
- $[\lambda x_1 \ldots x_n\, F^n x_1 \ldots x_n] = F^n$, for $n \geq 0$ (48.3)
- $([\lambda x_1 \ldots x_n\, \varphi]\!\downarrow \,\&\, \Box\forall x_1 \ldots \forall x_n(\varphi \equiv \psi)) \;\to\; [\lambda x_1 \ldots x_n\, \psi]\!\downarrow$ $(n \geq 1)$ (49)

Axioms for Encoding:

- $x_1 \ldots x_n F^n \;\equiv$ $x_1[\lambda y\, F^n y x_2 \ldots x_n] \,\&\, x_2[\lambda y\, F^n x_1 y x_3 \ldots x_n] \,\&\, \ldots \,\&\, x_n[\lambda y\, F^n x_1 \ldots x_{n-1} y]$ (50)
- $xF \to \Box xF$ (51)
- $O!x \to \neg\exists F\, xF$ (52)
- $\exists x(A!x \,\&\, \forall F(xF \equiv \varphi))$, provided x doesn't occur free in φ (53)

8.11 Explanatory Remarks: Digression

(**55**) **Remark:** Examples That Increase Understanding of Axiom (39.2). The following considerations may help one to better understand axiom (39.2). The axiom asserts:

> $\tau\downarrow$, whenever τ is either a primitive constant (i.e., one not introduced by a definition), a variable, or a core λ-expression.

In (39.2), we discussed why this axiom asserts that these terms are significant. But it may be helpful to give some examples of λ-expressions that aren't asserted to be significant by (39.2), i.e., non-core λ-expressions, in which the λ binds a variable in encoding position in the matrix (9.1). Consider the following extended, bulleted list of examples of such λ-expressions, some of which have been discussed before.

• (39.2) does not assert $[\lambda x\, \exists G(xG \,\&\, \neg Gx)]\downarrow$. In this expression, the λ binds the variable x, which occurs in encoding position (in the formula xG) in the matrix. This λ-expression will be provably empty (192), for otherwise one can derive the Clark/Boolos paradox.

• Consider the following definition:

$$\mathit{Clark\text{-}Boolos}(x) \equiv_{df} \exists G(xG \,\&\, \neg Gx)$$

Then (39.2) does not assert $[\lambda x\, \mathit{Clark\text{-}Boolos}(x)]\downarrow$. By Convention (17.3), x occurs in encoding position in $\mathit{Clark\text{-}Boolos}(x)$ since x occurs in encoding position in the definiens. Thus, $[\lambda x\, \mathit{Clark\text{-}Boolos}(x)]$ fails to be a core λ-expression and so violates the condition required by (39.2).

• Consider the following definition:

$$S_{10} =_{df} [\lambda x\, \exists G(xG \,\&\, \neg Gx)]$$

Then (39.2) does not assert $S_{10}\downarrow$, since it is a relation constant introduced by definition and not a primitive constant. The inferential role of this definition-by-= will ensure that $\neg S_{10}\downarrow$ follows from the fact that $\neg[\lambda x\, \exists G(xG \,\&\, \neg Gx)]\downarrow$, i.e., follows from theorem (192.1).

But also note that $[\lambda x\, S_{10}x]\downarrow$ is asserted by (39.2), since it is a core λ-expression! The variable bound by the λ doesn't occur in encoding position in the matrix $S_{10}x$. Moreover, the definiens of S_{10} doesn't contain any free occurrences of x, and our Encoding Formula Convention (17.3) doesn't 'kick in'. So the only variable bound by the λ in $[\lambda x\, S_{10}x]$, namely x, occurs in exemplification position. Although (39.2) rules that $[\lambda x\, S_{10}x]$ is significant, the property denoted will be provably empty, for given that such a property exists, β-Conversion yields, for an arbitrary x, that $[\lambda x\, S_{10}x]x \equiv S_{10}x$. Since we explained above why $\neg S_{10}\downarrow$, it follows from axiom (39.5.a) that $\neg S_{10}x$. Hence $\neg[\lambda x\, S_{10}x]x$, and this holds for any x.

• (39.2) does not assert $[\lambda x\, G\iota y(y=x \,\&\, \exists H(xH \,\&\, \neg Hx))]\downarrow$. The λ-expression is not a core λ-expression; the λ binds a variable that occurs in encoding position in two places, namely, in $y=x$ and in xH. This λ-expression will be provably empty, on pain of the Kirchner paradox.

• Consider the following definition:

$$c_x =_{df} \iota y(y=x \,\&\, \exists H(xH \,\&\, \neg Hx))$$

Now consider the λ-expression $[\lambda x\, Gc_x]$. (39.2) does not assert $[\lambda x\, Gc_x]\downarrow$. Note that if (39.2) were to assert $[\lambda x\, Gc_x]\downarrow$, then we could derive a contradiction.[122] By not asserting $[\lambda x\, Gc_x]\downarrow$ (and its closures), (39.2) avoids this result. In the definiens of c_x, the x occurs in encoding position (in two places). So by convention, the occurrence of x in c_x also occurs in encoding position. Thus, the λ in $[\lambda x\, Gc_x]$ binds a variable that occurs in encoding position and so fails to be a core λ-expression.

• To construct our last example of a non-core λ-expression that fails to meet the conditions of (39.2), we first define a core λ-expression. Let c_x be defined as above and let P again be any universal property. Then $[\lambda y\, Pc_x]$ is a core λ-expression (the y bound by the λ doesn't occur in encoding position in Pc_x) and so (39.2) asserts $[\lambda y\, Pc_x]\downarrow$. Now since $[\lambda y\, Pc_x]\downarrow$, consider the definition:

$$Q_x =_{df} [\lambda y\, Pc_x]$$

[122] For by (39.2), the universal closure would also be axiom:

(A) $\forall G([\lambda x\, Gc_x]\downarrow)$

Now it will be a theorem (193.2)★ that:

(B) $\exists G(\neg[\lambda x\, G\iota y(y=x \,\&\, \exists H(xH \,\&\, \neg Hx))]\downarrow)$

(A) and (B) are inconsistent. To see why, note that every universal property, i.e., every property that is exemplified by every individual, is a witness to (B) – see the proof of (193.1)★. Since it is easy to show that there are universal properties, let P be such a property and instantiate P into (A), to obtain $[\lambda x\, Pc_x]\downarrow$. From this and the relevant instance of β-Conversion (48.2) we obtain:

$$[\lambda x\, Pc_x]x \equiv Pc_x$$

Now by the reasoning in the proof of (193.1)★, it can be shown that Pc_x is equivalent to $\exists H(xH \,\&\, \neg Hx)$. Hence it would follow that:

$$[\lambda x\, Pc_x]x \equiv \exists H(xH \,\&\, \neg Hx)$$

Since we've made no assumptions about the free variable x, the last line holds for all x:

$$\forall x([\lambda x\, Pc_x]x \equiv \exists H(xH \,\&\, \neg Hx))$$

And since we derived $[\lambda x\, Pc_x]\downarrow$ from (A), it follows by ∃I that:

$$\exists F\forall x(Fx \equiv \exists H(xH \,\&\, \neg Hx))$$

This is a problematic assertion that leads to a version of the Clark/Boolos paradox.

and then consider the expression $[\lambda x\, Q_x a]$. (39.2) does *not* assert $[\lambda x\, Q_x a]\downarrow$. Since x occurs free in Pc_x in encoding position, convention requires us to suppose that x occurs in encoding position in Q_x. Thus, $[\lambda x\, Q_x a]$ is not a core λ-expression.

(**56**) **Remark:** Digression on Alternatives to (45.4). Various alternatives axioms were considered before settling on the version asserted as axiom (45.4). It may be of interest to some readers to learn why these alternatives were ultimately rejected in favor of (45.4). The alternatives that we considered are:

(.1) $\exists x(E!x \,\&\, \Diamond\neg E!x)$

(.2) $\Diamond\exists x(E!x \,\&\, \Diamond\neg E!x) \,\&\, \Diamond\neg\exists x(E!x \,\&\, \Diamond\neg E!x)$

(.3) $\Diamond\exists x(E!x \,\&\, \Diamond\neg E!x)$

(.4) $\Diamond\exists x E!x \,\&\, \Diamond\neg\exists x E!x$

Note that none of these asserts exactly what (45.4) asserts, namely, that it is possible that there exists a concrete-but-not-actually-concrete object.

The decision for using (45.4) can be summarized as follows. By asserting the possible existence of a concrete-but-not actually-concrete object, axiom (45.4), or a theorem that it implies, play a central role:

- in the proof of the existence of at least two contingent properties (205.5);
- in the proof of the *existence* of at least two contingent propositions (211.4), and thus plays a role in the proof of the existence of at least four propositions (212.4);
- in *identifying* two contingent propositions (215)★, one of which is contingently false (i.e., false but possibly true) and the other of which is contingently true (i.e., true but possibly false);
- in the proof of (theorems used in the proof of) the possible existence of concrete objects (205.3) and in the necessary existence of ordinary objects (227.1);
- in the proof of the existence of at least two possible worlds (547.4);
- in the proof the *predecessor* relation $\mathbb{P}$ is not an empty relation, i.e., in the proof that $\exists x\exists y\mathbb{P}xy$ (803.1); and
- in the proof that natural cardinals are discernible (803.3).

This axiom accomplishes the above without requiring the existence of concrete objects and without requiring the existence of a possible world that is devoid of (contingently) concrete objects.

The explanation as to why none of the alternatives have all these features will be conducted informally, by invoking as yet undefined notions of *possible world* and *truth at* at a possible world, as well as presupposing some basic model and set theory. So it should be remembered: (a) that, ultimately, talk about possible worlds and truth at a possible world is to be given the analysis in Chapter 12, Section 12.2; (b) that object theory doesn't assume or require any set theory or model theory; and (c) that some notions and axioms of set-theory will be reconstructed in object theory in Chapter 10, Section 10.3.

(.1) asserts that there exists an object that is both concrete but possibly not, i.e., there exists a contingently concrete object. If one were to assume that there is such an object, say a, then we would know $E!a \,\&\, \Diamond\neg E!a$. Since the expression $E!a$ denotes a proposition, we could conclude $\exists p(p \,\&\, \Diamond\neg p)$, and that conclusion would remain once we discharge the assumption about a. Thus, (.1) allows us to prove *that there is* a contingently true proposition, though we couldn't actually express such a proposition in our language. And once we have established that there is a contingently true proposition, it is straightforward to show that there is a contingently false one; just take the negation of any witness to the existence of a contingent truth.

Despite these virtues of (.1), the reason for not adopting it should be clear: we cannot justifiably assert (.1) *a priori* given that it implies $\exists xE!x$, i.e., there exist concrete objects. If we could justifiably assert (.1) *a priori*, Berkeley's idealism would be convincingly refutable *a priori*. It isn't. *A posteriori* evidence derived from our senses leads us to assign $\exists xE!x$ a very high probability. Since we can't justifiably assert $\exists xE!x$ *a priori*, we can't assert (.1) *a priori*. By contrast, (45.4) can be asserted *a priori*; it is true *a priori* that there might be a concrete-but-not-actually-concrete object, though I guess that a strict Humean might deny this.

(.2) asserts that there might be contingently concrete objects and that there might not be. (.2) immediately yields the existence of a contingent proposition, namely, $\exists x(E!x \,\&\, \Diamond\neg E!x)$, since it asserts both that this proposition is possible and that its negation is possible. So if $\exists x(E!x \,\&\, \Diamond\neg E!x)$ is true, it is contingently true (since it is possibly false), and if false, it is contingently false (since it is possibly true). Of course, (.2) doesn't tell you whether $\exists x(E!x \,\&\, \Diamond\neg E!x)$ is contingently true or contingently false. But (.2) also has the nice feature that it grounds a presupposition of Leibniz's famous question, "Why is there something rather than nothing?" (Article 7, Principles of Nature and Grace, 1714, PW 199, G.vi 602). Leibniz poses this question in the context of considering the natural world, and it is not unreasonable to suppose that he is asking, "Why is there something contingently concrete rather than nothing contingently concrete?" This presupposes that there might be no contingently concrete objects, which is precisely what the second conjunct of (.2) asserts *a priori*.

However, this second conjunct of (.2) does raise some questions. In the idiom of possible worlds, (.2) implies that there exists a possible world in which there are no contingently concrete objects. But consider this counterargument: if the key to naturalizing abstract objects is to suppose that they arise from (or even regard them as) patterns of contingently concrete objects in the natural world, then if *there had been* no contingently concrete objects, *there would have been* no abstract objects, which seems to contradict the idea that abstract objects are necessary beings. Even if such an argument is cogent, it isn't knockdown; as long as the actual world in fact has contingently concrete objects and abstract objects arise from (or arise as) patterns of such objects, then it may not matter that a possible world empty of contingently concrete objects fails to give rise to abstract objects. So, although the second conjunct of (.2) gives rise to (the appearance of) controversy. Thus, we shall not assert (.2) given that the existence of a possible world devoid of contingently concrete objects may yield more metaphysical questions than it answers.

By contrast, (45.4) doesn't require the existence of a possible world devoid of contingently concrete objects. (45.4) is true in models in which there are exactly two possible worlds (the actual world $\boldsymbol{w}_\alpha$ and a distinct world $\boldsymbol{w}_1$) and two distinct objects, one of which is concrete at $\boldsymbol{w}_\alpha$ but not at $\boldsymbol{w}_1$ and the other of which is concrete at $\boldsymbol{w}_1$ but not at $\boldsymbol{w}_\alpha$. But (.2) rules out such models. Of course, (45.4) *allows* for models in which nothing is contingently concrete at $\boldsymbol{w}_\alpha$ as long as there is a world like $\boldsymbol{w}_1$ at which some object is concrete there but not at $\boldsymbol{w}_\alpha$. (Of course, it is unlikely that such a model represents the way things are.) So, in this sense, (45.4) is weaker than (.2). But (45.4) is stronger in the sense that it rules out models with exactly two possible worlds in which there is one contingently concrete object at $\boldsymbol{w}_\alpha$ and none at $\boldsymbol{w}_1$. Such a model is permitted by (.2).[123]

(.3) doesn't have the problem that (.2) has, since it simply drops the second conjunct of (.2). Though (.3) doesn't immediately yield the existence of a contingent proposition, we may derive the existence of such a proposition from (.3) *without* appealing to the current (45.4).[124] The disadvantages of (.3) emerge only in comparison with (45.4). From the latter, we shall not only

[123]By contrast, (.2) rules out, while (45.4) is true in, a model containing exactly two possible worlds, in which the actual world has a single contingently concrete object while the second world has two, since (.2) requires the existence of world *empty* of contingently concrete objects.

[124]Assume $\Diamond\exists x(E!x \,\&\, \Diamond\neg E!x)$. Then by BF$\Diamond$ (167.3), it follows that $\exists x\Diamond(E!x \,\&\, \Diamond\neg E!x)$. Suppose a is such an object, so that we know $\Diamond(E!a \,\&\, \Diamond\neg E!a)$. Since a possible conjunction implies that the conjuncts are possible, it follows that $\Diamond E!a$ and $\Diamond\Diamond\neg E!a$. But by a relevant instance of the 4$\Diamond$ schema (165.7), the latter implies $\Diamond\neg E!a$. So we have established $\Diamond E!a \,\&\, \Diamond\Diamond\neg E!a$. Now we also know that $[\lambda x\, E!a]\downarrow$, by (39.2), and so it follows that $(E!a)\downarrow$, by definition (20.3). Since the formula $E!a$ has a denotation, $\Diamond E!a \,\&\, \Diamond\Diamond\neg E!a$ implies $\exists p(\Diamond p \,\&\, \Diamond\neg p)$. Thus, we have derived that there exists a contingent proposition. This proof, however, is non-constructive; we can't express a witness in our language.

be able to prove that there exist contingent propositions, but we'll be able to identify a particular proposition that is contingently true and one that is contingently false. We'll see in (215.1)★ that $\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$ is contingently false, and in (215.2)★ that its negation is contingently true. By contrast, (.3) isn't strong enough to yield these results.

Finally, (.4) has the virtue that it, too, implies the existence of a contingent proposition, namely, $\exists xE!x$, since it asserts both that this proposition and its negation are both possible. So if $\exists xE!x$ is true, it is contingently true (since it is possibly false), and if false, contingently false (since it is possibly true). But there are two problems with the second conjunct of (.4). Since the second conjunct requires that there be a world devoid of *any* concrete objects, it requires both (a) the existence of a world where there are no contingently concrete objects (i.e., objects that are concrete there but not concrete with respect to some other world) and (b) the existence of a world where there are no *necessarily concrete* objects (like Spinoza's God). We've already discussed the questions that arise in connection with (a). As for (b), if Spinoza is correct that God just is Nature and that God is a necessary being, then given that Nature is concrete, it would follow that God (g) is necessarily concrete, i.e., that $\Box E!g$. If so, it wouldn't be correct to assert that it is possible that there are no concrete objects; at least, we shouldn't assert this *a priori*. Note that in giving this argument, we're not *agreeing* with Spinoza's view; only suggesting that object theory should not rule his view out *a priori*. (Of course, some may think it would be a feature of the system if it were to rule out Spinoza's view.) In any case, (45.4) doesn't have the disadvantages of (.4).

(**57**) **Remark:** Digression on the Coexistence of Relations. It may be of some interest to learn why the 2nd conjunct of the antecedent of (49) must include both a modal operator and universal quantifiers.[125] Consider what the unary case of (49) would have asserted if the modal operator and universal quantifier had been removed from the second conjunct:

$$(\vartheta)\quad ([\lambda x\, \varphi]\!\downarrow \,\&\, (\varphi \equiv \psi)) \rightarrow [\lambda x\, \psi]\!\downarrow$$

To see why the universal quantifier $\forall x$ must preface the second conjunct of the antecedent, we derive a contradiction from (ϑ) that would have been avoidable had the quantifier been present. Let φ be $\forall p(p \rightarrow p)$ and let ψ be the formula that leads to the Kirchner paradox, i.e., $G\iota y(y = x \,\&\, \exists H(xH \,\&\, \neg Hx))$, where G is some universal property such as $[\lambda z\, \forall p(p \rightarrow p)]$. Then the following is an instance of (ϑ):

[125]The argument that follows was sketched by Daniel Kirchner, in personal communications (email correspondence) of 17 November 2017 and thereafter. For interested readers, we have reconstructed the argument in greater detail, by considering the unary case of the axiom.

$$(\xi)\ \ ([\lambda x\, \forall p(p \to p)]{\downarrow}\ \&\ (\forall p(p \to p) \equiv G\iota y(y = x\ \&\ \exists H(xH\ \&\ \neg Hx)))) \to [\lambda x\, G\iota y(y = x\ \&\ \exists H(xH\ \&\ \neg Hx))]{\downarrow}$$

Note that the variable x has three free occurrences in the second conjunct of the antecedent. These occurrences become bound in the universal closure of (ξ), which would, like (ξ), also be an axiom if (ϑ) were an axiom:

$$(\zeta)\ \ \forall x(([\lambda x\, \forall p(p \to p)]{\downarrow}\ \&\ (\forall p(p \to p) \equiv G\iota y(y = x\ \&\ \exists H(xH\ \&\ \neg Hx)))) \to [\lambda x\, G\iota y(y = x\ \&\ \exists H(xH\ \&\ \neg Hx))]{\downarrow})$$

Now consider any abstract object that encodes $O!$; for example, let a be an abstract object that encodes just $O!$.[126] So a is such that $A!a$ and $\forall F(aF \equiv F = O!)$. Then we know not only that $aO!$, but also, since $A!a$, that $\neg O!a$, by (222.3). Hence, $\exists H(aH\ \&\ \neg Ha)$. Now if we instantiate a into (ζ), we obtain:

$$(\xi')\ \ ([\lambda x\, \forall p(p \to p)]{\downarrow}\ \&\ (\forall p(p \to p) \equiv G\iota y(y = a\ \&\ \exists H(aH\ \&\ \neg Ha)))) \to [\lambda x\, G\iota y(y = x\ \&\ \exists H(xH\ \&\ \neg Hx))]{\downarrow}$$

But both conjuncts of the antecedent of (ξ') would now be derivable. The first conjunct is an axiom, since the λ-expression meets the conditions of (39.2). It then remains only to show:

$$\forall p(p \to p) \equiv G\iota y(y = a\ \&\ \exists H(aH\ \&\ \neg Ha))$$

Proof. Since $\forall p(p \to p)$ is a theorem, it suffices to show only that the right condition is a theorem. But when p is true, $\iota y(y = x)$ is identical to $\iota y(y = x\ \&\ p)$. Moreover both descriptions have the same denotation as x. Since we've established $\exists H(aH\ \&\ \neg Ha)$, it follows that $\iota y(y = a\ \&\ \exists H(aH\ \&\ \neg Ha))$ denotes a. So we only need to show Ga. But this follows from the assumption that G is a universal property. ⋈

So it would follow from (ξ') that $[\lambda x\, G\iota y(y = x\ \&\ \exists H(xH\ \&\ \neg Hx))]{\downarrow}$, i.e., that the λ-expression has a denotation. Since we know this result leads to a contradiction, we've shown why (ϑ) has to be replaced by:

$$(\omega)\ \ ([\lambda x\, \varphi]{\downarrow}\ \&\ \forall x(\varphi \equiv \psi)) \to [\lambda x\, \psi]{\downarrow}$$

With the quantifier $\forall x$ prefacing the second conjunct of the antecedent, the axiom can't be true simply on the basis of a single assignment to the free variable x in (ξ).

But (ω) still won't do as an axiom. We must add not only the quantifier $\forall x$ to the second conjunct of the antecedent, but also a modal operator, on pain of contradiction. To see how the contradiction would arise, we have

[126] The claim $\exists x(A!x\&\forall F(xF \equiv F = O!))$ is an instance of the Comprehension Principle for Abstract Objects (53). It asserts asserts the existence of an abstract object that encodes just $O!$ and no other properties.

to appeal to modal theorem (215.2)★, which identifies a particular, contingently true proposition.[127] To state this theorem, we let q_0 be the proposition $\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$. Then (215.2)★ asserts that the negation of q_0 is contingently true, i.e., that $ContingentlyTrue(\overline{q_0})$. So by definition (213.1), we know $\overline{q_0} \& \Diamond\neg\overline{q_0}$, i.e.,

(A) $\neg q_0 \,\&\, \Diamond q_0$

Then we derive a contradiction from this fact and (ω) as follows.

Let χ be the matrix of the expression used in the Kirchner's Paradox, i.e., let χ be $G\imath y(y = x \,\&\, \exists H(xH \,\&\, \neg Hx))$. (As we saw earlier in (55), there are two occurrences of x in χ that are in encoding position, one of which is buried in the identity claim.) Since we know $\neg q_0$ by (A), we also know $\neg q_0 \vee \mathcal{A}\chi$. Hence we know the following equivalence, since the left side is easily derivable and the right side has been shown to be true:

$$\forall p(p \to p) \equiv (\neg q_0 \vee \mathcal{A}\chi)$$

Hence by GEN:

(B) $\forall x(\forall p(p \to p) \equiv (\neg q_0 \vee \mathcal{A}\chi))$

Since $[\lambda x\, \forall p(p \to p)]\downarrow$ (39.2), it follows from (ω) that:

(C) $[\lambda x\, \neg q_0 \vee \mathcal{A}\chi]\downarrow$

Now the following is an axiom, since it is the universal generalization of an instance of (39.2):

(D) $\forall G([\lambda x\, \Box Gx]\downarrow)$

So from (C) and (D) it follows that:

(E) $[\lambda x\, \Box[\lambda x\, \neg q_0 \vee \mathcal{A}\chi]x]\downarrow$

From (E) and an instance of β-Conversion (48.2) it follows that:

(F) $[\lambda x\, \Box[\lambda x\, \neg q_0 \vee \mathcal{A}\chi]x]x \equiv \Box[\lambda x\, \neg q_0 \vee \mathcal{A}\chi]x$

Now theorem (106) tells us that if anything exists in the logical sense, it does so necessarily, i.e., that $\tau\downarrow \to \Box\tau\downarrow$. So, it follows from (C) that:

[127] Since the following, extended reasoning appeals both to theorem (215.2)★ and theorem (130.2)★, it is not *modally strict* in the sense of (60.2). But a contradiction is to be avoided, no matter whether it arises by modally strict or non-modally strict reasoning. We might be able to derive a contradiction by *modally strict* means if we start with theorem (217.1), which asserts that there are contingently true propositions. But (a) the present reasoning is simpler if we just identify a specific, contingently true proposition, and (b) it suffices to make the point by deriving a contradiction by non-modally strict reasoning.

(G) $\Box[\lambda x\, \neg q_0 \vee \mathcal{A}\chi]\downarrow$

Now the following is axiomatic, since it is the modal closure of an instance of β-Conversion:

(H) $\Box([\lambda x\, \neg q_0 \vee \mathcal{A}\chi]\downarrow \rightarrow ([\lambda x\, \neg q_0 \vee \mathcal{A}\chi]x \equiv (\neg q_0 \vee \mathcal{A}\chi)))$

Hence, by the K axiom (45.1), it follows from (G) and (H) that:

(I) $\Box([\lambda x\, \neg q_0 \vee \mathcal{A}\chi]x \equiv (\neg q_0 \vee \mathcal{A}\chi))$

Now (I) implies by the modal theorem (158.6) that:

(J) $\Box[\lambda x\, \neg q_0 \vee \mathcal{A}\chi]x \equiv \Box(\neg q_0 \vee \mathcal{A}\chi)$

Hence, by biconditional syllogism, (F) and (J) imply:

(K) $[\lambda x\, \Box[\lambda x\, \neg q_0 \vee \mathcal{A}\chi]x]x \equiv \Box(\neg q_0 \vee \mathcal{A}\chi)$

Put (K) aside for the moment and let's reason about its right side. Note that given the second conjunct of (A), we can establish:

(L) $\Box(\neg q_0 \vee \mathcal{A}\chi) \equiv \Box\mathcal{A}\chi$

> *Proof*. ($\rightarrow$) Assume $\Box(\neg q_0 \vee \mathcal{A}\chi)$. This assumption and $\Diamond\neg\neg q_0$ (which we can easily obtain from the second conjunct of (A), i.e., from $\Diamond q_0$) imply, by (162.7), that $\Diamond\mathcal{A}\chi$. But actuality claims are subject to modal collapse; i.e., by theorem (174.1), we know $\Diamond\mathcal{A}\varphi \equiv \Box\mathcal{A}\varphi$. Hence, $\Box\mathcal{A}\chi$. ($\leftarrow$) Assume $\Box\mathcal{A}\chi$. *A fortiori*, $\Box(\neg q_0 \vee \mathcal{A}\chi)$.

So, (K) and (L) imply:

(M) $[\lambda x\, \Box[\lambda x\, \neg q_0 \vee \mathcal{A}\chi]x]x \equiv \Box\mathcal{A}\chi$

Put (M) aside for the moment and reason about its right side. Note that we can establish:

(N) $\Box\mathcal{A}\chi \equiv \chi$

> *Proof*. As an instance of axiom (46.1), we know $\mathcal{A}\chi \rightarrow \Box\mathcal{A}\chi$. But the T schema yields $\Box\mathcal{A}\chi \rightarrow \mathcal{A}\chi$. By conjoining the latter and former and applying the definition of $\equiv$, we obtain $\Box\mathcal{A}\chi \equiv \mathcal{A}\chi$. But a non-modally strict theorem governing the logic of actuality (130.2)$\star$ tells us $\mathcal{A}\chi \equiv \chi$. Hence by biconditional syllogism, $\Box\mathcal{A}\chi \equiv \chi$.

So from (M) and (N) we have:

$$[\lambda x\, \Box[\lambda x\, \neg q_0 \vee \mathcal{A}\chi]x]x \equiv \chi$$

Since we've derived this from no assumptions, it follows by GEN that:

$$\forall x([\lambda x\, \Box[\lambda x\, \neg q_0 \vee \mathcal{A}\chi]x]x \equiv \chi)$$

Since we've already established that $[\lambda x\, \Box[\lambda x\, \neg q_0 \vee \mathcal{A}\chi]x]\downarrow$, it follows by $\exists$I (101) that:

$$\exists F \forall x(Fx \equiv \chi)$$

But recall that χ is the matrix of the expression used in the Kirchner paradox, which means we have established:

$$\exists F \forall x(Fx \equiv G\imath y(y = x \,\&\, \exists H(xH \,\&\, \neg Hx)))$$

And this, we know, leads to a contradiction when G is a necessary property.

We avoid this result by prefacing the modal operator $\Box$ to the second conjunct of the antecedent of (ω). By doing so, the conclusion (C) only follows if we can show that (B) is necessary. But (B) can't be necessary if $\neg q_0$ is contingent. Intuitively, consider any possible world, say $\boldsymbol{w}$, where both $\neg q_0$ is false (i.e., where q_0 is true) and where there is an object x that fails to be such that $\mathcal{A}\chi$. Then (B) is false at $\boldsymbol{w}$: something is such that the left side of (B) is true at $\boldsymbol{w}$ while the right side of (B) is false at $\boldsymbol{w}$. Thus, the second conjunct of the antecedent of axiom (49) has to be a necessarily true universal claim.

Chapter 9

Deductive Systems of PLM

In science, what is provable shouldn't be believed without proof.[128]

— Dedekind 1888

In this chapter, we introduce the deductive system PLM by combining the axioms of the previous chapter with a primitive rule of inference and defining notions of proof and theoremhood. We then develop a series of basic theorems, facilitated by the introduction and justification of metarules that help us to establish theorems more easily. Many of the proofs of theorems and justifications of metarules are left to the main Appendix, though some metatheorems are proved in chapter appendices.

Readers are reminded that object-language variables function normally in axioms and theorems, but function as metavariables when they appear in definitions, as per Convention (17.2).

9.1 Primitive Rule of PLM: Modus Ponens

(**58**) **Primitive Rule of Inference**: Modus Ponens. PLM employs just a single primitive rule of inference:

Modus Ponens (**Rule MP**)
$\varphi, \varphi \to \psi \; / \; \psi$

We may read this as: ψ follows from φ and $\varphi \to \psi$.

[128] Translation mine. The original German is:

> Was beweisbar ist, soll in der Wissenschaft nicht ohne Beweiß geglaubt werden.

This is the opening line from the Preface to Dedekind's classic work of 1888; it occurs in the second edition (1893) on p. vii and in the seventh edition (1939) on p. iii.

9.2 (Modally Strict) Proofs and Derivations

(**59**) **Metadefinitions:** Derivations, Proofs, and Theorems of PLM. In what follows, we say that φ is an *axiom* of PLM whenever φ is one of the axioms asserted in Chapter 8. The set of axioms of PLM is recursive and we introduce the following symbol to refer to it:

$$\boldsymbol{\Lambda} = \{\,\varphi \,|\, \varphi \textit{ is an axiom}\,\}$$

Then we define:

(.1) A *derivation* in PLM *of* φ *from* a set of formulas Γ is any non-empty sequence of formulas $\varphi_1,\ldots,\varphi_n$ such that φ is φ_n and for each i such that $1 \leq i \leq n$, φ_i either is an element of $\boldsymbol{\Lambda} \cup \Gamma$ or follows from two of the preceding members of the sequence by the Rule MP. When there exists a derivation of φ from some Γ, we write $\Gamma \vdash \varphi$ and we say either φ *is derivable from* Γ, or φ *is a logical consequence of* Γ (in the proof-theoretic sense), or more simply φ *follows from* Γ.

Thus, $\Gamma \vdash \varphi$ expresses a multigrade metatheoretical relation between zero or more formulas in Γ (traditionally called the *premises* or *assumptions*) and φ (the *conclusion*).

The following conventions apply. We often write $\varphi_1,\ldots,\varphi_n \vdash \psi$ when $\Gamma \vdash \psi$ and Γ is the set $\{\varphi_1,\ldots,\varphi_n\}$. We often write $\Gamma,\psi \vdash \varphi$ when $\Gamma \cup \{\psi\} \vdash \varphi$. We often write $\Gamma_1,\Gamma_2 \vdash \varphi$ when $\Gamma_1 \cup \Gamma_2 \vdash \varphi$.[129]

When a sequence of formulas ending in φ is a derivation of φ from some set Γ, we say the sequence is a *witness to* (the claim that) $\Gamma \vdash \varphi$ and we call the members of the sequence the *lines* of the derivation.

Now using the definition of *derivation*, we may define the notions of *proof* in PLM and *theorem* of PLM:

(.2) A *proof* of φ in PLM is any derivation of φ from Γ in PLM in which Γ is the empty set $\varnothing$. A formula φ is a *theorem* of PLM, written $\vdash \varphi$, if and only if there exists a proof of φ in PLM.

Two simple consequences of our definitions are:

(.3) $\vdash \varphi$ if and only if there is a non-empty sequence of formulas $\varphi_1,\ldots,\varphi_n$ such that φ is φ_n and for each i $(1 \leq i \leq n)$, φ_i either is an element of $\boldsymbol{\Lambda}$ or follows from two of the preceding members of the sequence by the Rule MP.

[129]Consequently, since $\Gamma_1 \cup \Gamma_2 = \Gamma_2 \cup \Gamma_1$, the order in which premise sets are listed doesn't matter, and so in the case where Γ_1 and Γ_2 are singletons, the order in which the premises are listed doesn't matter. Thus, in general, $\Gamma_1,\Gamma_2 \vdash \varphi$ if and only if $\Gamma_2,\Gamma_1 \vdash \varphi$, and in particular, $\varphi,\psi \vdash \chi$ if and only if $\psi,\varphi \vdash \chi$. See (63.11).

(.4) If $\Gamma = \varnothing$, then $\Gamma \vdash \varphi$ if and only if $\vdash \varphi$.

(**60**) **Metadefinitions:** Modally Strict Proofs, Theorems, and Derivations. To precisely identify those derivations and proofs in which no inferential step is justified by a modally fragile axiom, we first say that φ is a *necessary* axiom whenever φ is any axiom such that *all* the closures of φ are also axioms. At present, (43)★ is the only axiom that fails to be necessary in this sense.

We introduce the following symbol to refer to the set of necessary axioms:

$$\mathbf{\Lambda}_\Box = \{\varphi \,|\, \varphi \textit{ is a necessary axiom}\}$$

Then the following definitions of modally-strict derivations ($\Gamma \vdash_\Box \varphi$) and modally-strict proofs ($\vdash_\Box \varphi$) mirror (59.1) and (59.2), with the exception that the definientia refer to $\mathbf{\Lambda}_\Box$ instead of $\mathbf{\Lambda}$:

(.1) A *modally-strict derivation* (or $\Box$-*derivation*) *of* φ *from* a set of formulas Γ in PLM is any non-empty sequence of formulas $\varphi_1, \ldots, \varphi_n$ such that $\varphi = \varphi_n$ and for each i ($1 \le i \le n$), φ_i is either an element of $\mathbf{\Lambda}_\Box \cup \Gamma$ or follows from two of the preceding members of the sequence by Rule MP. A formula φ *is strictly derivable* (or $\Box$-derivable) *from* the set Γ in PLM, written $\Gamma \vdash_\Box \varphi$, just in case there exists a modally-strict derivation of φ from Γ.

(.2) A *modally-strict proof* (or $\Box$-*proof*) of φ in PLM is any modally-strict derivation of φ from Γ when Γ is the empty set. A formula φ is a *modally-strict theorem* (or $\Box$-*theorem*) of PLM, written $\vdash_\Box \varphi$, if and only if there exists a modally-strict proof of φ in PLM.

These two definitions have simple consequences analogous to (59.3) and (59.4). We shall suppose that all of the conventions introduced in (59) concerning $\vdash$ also apply to $\vdash_\Box$.

(**61**) **Remark:** Metarules of Inference. In what follows, we often introduce, and sometimes prove, certain claims about derivations. These facts all have the following form:

If conditions … hold, then there exists a derivation of φ from Γ.

When call facts of this form *metarules of inference*. Only two metarules of inference are taken as primitive and underived, namely, the Rule of Definition by Equivalence (72) and the Rule of Definition by Identity (73). The other metarules are proved or left as exercises in the Appendix.

Metarules of inference are to be contrasted with standard *rules of inference*; the latter allow us to infer φ from zero or more formulas, whereas the former allow us to infer the existence of a derivation or proof of φ given certain conditions. In what follows, when we are deriving φ from Γ, we can stop the reasoning when we've met the conditions of a metarule whose consequent asserts that there is a sequence of formulas constituting a derivation of φ from Γ. So metarules often shorten the reasoning we use in the Appendix to establish that $\Gamma \vdash \varphi$.

Consequently, when we reason with metarules to establish the claim that $\Gamma \vdash \varphi$, we don't produce a witness to the claim. However, the proof of a derived metarule in the Appendix shows how to construct such a witness. As mentioned previously, we call the proof of a derived metarule its *justification*. The justification shows that reasoning with a metarule can always be converted reasoning without it.

(**62**) **Metarules:** Modally Strict Derivations are Derivations. It immediately follows from our definitions that: (.1) if there is a modally-strict derivation of φ from Γ, then there is a derivation of φ from Γ, and (.2) if there is a modally-strict proof of φ, then there is a proof of φ:

(.1) If $\Gamma \vdash_\Box \varphi$, then $\Gamma \vdash \varphi$

(.2) If $\vdash_\Box \varphi$, then $\vdash \varphi$

Clearly, however, the converses are not true in general, since derivations and proofs in which the modally fragile axiom (43)★ is used are not modally strict. Consequently, modally-strict derivations and proofs constitute a proper subset, respectively, of all derivations and proofs.

(**63**) **Metarules:** Fundamental Properties of $\vdash$ and $\vdash_\Box$. The following facts are particularly useful as we prove new theorems and justify new metarules of PLM. Note that these facts come in pairs, with one member of the pair governing $\vdash$ and the other member governing $\vdash_\Box$:

(.1) If $\varphi \in \mathbf{\Lambda}$, then $\vdash \varphi$. ("Axioms are theorems")
If $\varphi \in \mathbf{\Lambda}_\Box$, then $\vdash_\Box \varphi$. ("Necessary axioms are modally-strict theorems")

(.2) If $\varphi \in \Gamma$, then $\Gamma \vdash \varphi$. (Note the special case: $\varphi \vdash \varphi$)
If $\varphi \in \Gamma$, then $\Gamma \vdash_\Box \varphi$. (Note the special case: $\varphi \vdash_\Box \varphi$)

(.3) If $\vdash \varphi$, then $\Gamma \vdash \varphi$.
If $\vdash_\Box \varphi$, then $\Gamma \vdash_\Box \varphi$.

(.4) If $\varphi \in \mathbf{\Lambda} \cup \Gamma$, then $\Gamma \vdash \varphi$.
If $\varphi \in \mathbf{\Lambda}_\Box \cup \Gamma$, then $\Gamma \vdash_\Box \varphi$.

(.5) If $\Gamma_1 \vdash \varphi$ and $\Gamma_2 \vdash (\varphi \to \psi)$, then $\Gamma_1, \Gamma_2 \vdash \psi$.
If $\Gamma_1 \vdash_\Box \varphi$ and $\Gamma_2 \vdash_\Box (\varphi \to \psi)$, then $\Gamma_1, \Gamma_2 \vdash_\Box \psi$.

(.6) If $\vdash \varphi$ and $\vdash (\varphi \to \psi)$, then $\vdash \psi$.
If $\vdash_\Box \varphi$ and $\vdash_\Box (\varphi \to \psi)$, then $\vdash_\Box \psi$.

(.7) If $\Gamma \vdash \varphi$ and $\Gamma \subseteq \Delta$, then $\Delta \vdash \varphi$.
If $\Gamma \vdash_{\Box} \varphi$ and $\Gamma \subseteq \Delta$, then $\Delta \vdash_{\Box} \varphi$.

(.8) If $\Gamma \vdash \varphi$ and $\varphi \vdash \psi$, then $\Gamma \vdash \psi$.
If $\Gamma \vdash_{\Box} \varphi$ and $\varphi \vdash_{\Box} \psi$, then $\Gamma \vdash_{\Box} \psi$.

(.9) If $\Gamma \vdash \varphi$, then $\Gamma \vdash (\psi \rightarrow \varphi)$, for any ψ.
If $\Gamma \vdash_{\Box} \varphi$, then $\Gamma \vdash_{\Box} (\psi \rightarrow \varphi)$, for any ψ.

(.10) If $\Gamma \vdash (\varphi \rightarrow \psi)$, then $\Gamma, \varphi \vdash \psi$.
If $\Gamma \vdash_{\Box} (\varphi \rightarrow \psi)$, then $\Gamma, \varphi \vdash_{\Box} \psi$.

(.11) If $\Gamma_1, \ldots, \Gamma_i, \ldots, \Gamma_j, \ldots, \Gamma_n \vdash \varphi$, then $\Gamma_1, \ldots, \Gamma_j, \ldots, \Gamma_i, \ldots, \Gamma_n \vdash \varphi$ $\quad (1 \leq i \leq j \leq n)$
If $\Gamma_1, \ldots, \Gamma_i, \ldots, \Gamma_j, \ldots, \Gamma_n \vdash_{\Box} \varphi$, then $\Gamma_1, \ldots, \Gamma_j, \ldots, \Gamma_i, \ldots, \Gamma_n \vdash_{\Box} \varphi$

Fact (.11) indicates that the order in which premise sets are listed makes no difference. Note that in the case where $\Gamma_1, \ldots, \Gamma_n$ are singletons, it follows that the order in which the *premises* are listed doesn't matter. Thus, it follows as a special case of (.11) that:

If $\psi_1, \ldots, \psi_i, \ldots, \psi_j, \ldots, \psi_n \vdash \varphi$, then $\psi_1, \ldots, \psi_j, \ldots, \psi_i, \ldots, \psi_n \vdash \varphi$ $\quad (1 \leq i \leq j \leq n)$
If $\psi_1, \ldots, \psi_i, \ldots, \psi_j, \ldots, \psi_n \vdash_{\Box} \varphi$, then $\psi_1, \ldots, \psi_j, \ldots, \psi_i, \ldots, \psi_n \vdash_{\Box} \varphi$

Notice also that in the special case of (.2), $\varphi \vdash_{\Box} \varphi$ holds even if φ isn't a necessary truth. In general, there can be modally-strict derivations in which neither the premises nor conclusion are necessary truths. The 3-element sequence $Pa, Pa \rightarrow Qb, Qb$ is a modally-strict derivation of Qb from the assumptions Pa and $Pa \rightarrow Qb$ whether or not the premises and the conclusion are necessary truths.

(**64**) **Remark:** Theorems That Aren't Modally Strict ($\star$-Theorems). For the most part, we shall be interested in proofs generally, not just modally-strict ones, since our primary interest is what claims we can prove (*simpliciter*). But since significantly more modally-strict $\Box$-theorems are proved in what follows, it is useful to mark the ones that are not. So we introduce the $\star$ annotation to make it explicit that a proof of a theorem is *not* modally-strict. The reader may therefore assume that all of the items marked **Theorem** in what follows have modally-strict proofs, and that $\star$-**Theorems** do not. We always concatenate a theorem's item number with a $\star$ when referencing a $\star$-theorem.[130] Similarly, in the case where the derivation of φ from Γ is not modally-strict, we may speak of $\star$-derivations and say that φ is $\star$-derivable from Γ.

It is important to recognize that $\star$-theorems and $\star$-derivations are not defective in any way. Indeed, they are simply artifacts of a modal logic that allows

[130] The first such theorems are (138.1)$\star$ – (138.2)$\star$ and (140)$\star$ – (145.4)$\star$ below.

one to reason from modally fragile axioms or contingent premises, which are often of great philosophical significance.

(**65**) **Metadefinition:** Dependence. It is sometimes useful to indicate the difference between $\square$-derivations and $\star$-derivations by saying that in the latter, the conclusion *depends* upon a modally fragile axiom, or depends upon a $\star$-theorem that in turn depends upon a modally fragile axiom, etc. To make this talk of dependence precise, we define the conditions under which one formula depends upon another within the context of a derivation:

> Let the sequence $\varphi_1, \ldots, \varphi_n$ be a derivation in PLM of φ (= φ_n) from the set of premises Γ and let ψ be a member of this sequence. Then we say that φ_i ($1 \leq i \leq n$) *depends upon* the formula ψ *in* this derivation iff either (a) $\varphi_i = \psi$, or (b) φ_i follows by the Rule MP from two previous members of the sequence at least one of which depends upon ψ.

When $\Gamma = \varnothing$, the definiendum in the above becomes: φ *depends on* a formula ψ in a given proof of φ. The $\star$-theorems formulated in what follows depend upon axiom (43)$\star$ either because some inferential step directly cites this axiom, or because some inferential step cites a $\star$-theorem that directly cites this axiom, etc.

It follows from our definition that if a sequence S is a witness to $\Gamma \vdash \varphi$, then S is a witness to $\Gamma \vdash_\square \varphi$ if and only if φ doesn't depend upon any instance of a modally fragile axiom in S. This holds even if, in S, φ depends upon a premise in Γ that isn't necessary. The sequence $S = Pa, Pa \rightarrow Qb, Qb$ is a witness to $Pa, Pa \rightarrow Qb \vdash Qb$ and even if Pa fails to be a necessary truth, it follows that $Pa, Pa \rightarrow Qb \vdash_\square Qb$, since Qb doesn't depend on (43)$\star$ in S. (This case will be discussed in some detail below, when we introduce the Rule of Necessitation.)

9.3 Two Fundamental Metarules: GEN and RN

(**66**) **Metarule**: The Rule of Universal Generalization. The Rule of Universal Generalization (GEN) asserts that whenever there is a derivation of φ from a set of premises Γ, and the variable α doesn't occur free in any of the premises in Γ, then there is a derivation from Γ of the claim $\forall\alpha\varphi$ (even if α occurs free in φ):

> **Rule of Universal Generalization** (**GEN**)
> If $\Gamma \vdash \varphi$ and α doesn't occur free in any formula in Γ, then $\Gamma \vdash \forall\alpha\varphi$.

When $\Gamma = \varnothing$, then GEN asserts that if a formula φ is a theorem, then so is $\forall\alpha\varphi$:

> If $\vdash \varphi$, then $\vdash \forall\alpha\varphi$

We note here that the justification of this rule in the Appendix can be easily converted to a justification of a $\vdash_\Box$ version of GEN, in which $\vdash_\Box$ is substituted everywhere for $\vdash$. We leave further discussion of this to Remark (67).

When Γ isn't empty, the application of GEN requires that the variable α not occur free in any premise in Γ. This prohibits one from using GEN, for example, to derive $\forall x Rx$ from the premise Rx. We know $Rx \vdash Rx$ by the special case of (63.2), but intuitively, $\forall x Rx$ doesn't follow from Rx; from the premise that a particular, but unspecified, individual x exemplifies the property R, it doesn't follow that every individual exemplifies R. The proviso to GEN, of course, is unnecessary when φ is a theorem since φ is then derivable from the empty set of premises. Whenever any formula φ with free variable α is a theorem, we may invoke GEN to conclude that $\forall\alpha\varphi$ is also a theorem. For example, we shall soon prove that $\varphi \to \varphi$ is a theorem (74), so that the instance $Px \to Px$ is a theorem. From this latter it follows by GEN that $\forall x(Px \to Px)$.

Here is an example of GEN in action. The following reasoning sequence establishes that $\forall x(Qx \to Px)$ is derivable from the premise $\forall x Px$, even though strictly speaking, the sequence is not a witness to this derivability claim:

1.	$\forall x Px$	Premise
2.	$\forall x Px \to (x{\downarrow} \to Px)$	Instance, Axiom (39.1)
3.	$x{\downarrow} \to Px$	from 1,2, by MP
4.	$x{\downarrow}$	Instance, Axiom (39.2)
5.	Px	from 3,4, by MP
6.	$Px \to (Qx \to Px)$	Instance, Axiom (38.1)
7.	$Qx \to Px$	from 5, 6, by MP
8.	$\forall x Px \vdash Qx \to Px$	from 1–7, by df $\Gamma \vdash \varphi$ (59)
9.	$\forall x Px \vdash \forall x(Qx \to Px)$	from 8, by GEN

Line 9 asserts that there is a derivation of $\forall x(Qx \to Px)$ from $\forall x Px$, but lines 1–8 are *not* such a derivation. However, lines 1–7 are a witness to line 8: each of lines 1–7 is either an axiom, a premise, or follows by MP from two previous lines. Since line 8 is a claim that has the form of the antecedent to the metarule GEN, and the condition, that x is not free in the premise $\forall x Px$, is met, we may apply GEN to obtain line 9.

Though lines 1–8 do not constitute a witness to $\forall x Px \vdash \forall x(Qx \to Px)$, the justification (i.e., metatheoretic proof) of GEN given in the Appendix shows us how to convert the reasoning into a sequence of formulas that is a *bona fide* witness to the derivability claim. By studying the metatheoretic proof, it becomes clear that the above reasoning with GEN can be converted to the following derivation, in which no appeal to GEN is made:[131]

[131] Line 2 in the following derivation is an axiom: $Px \to (Qx \to Px)$ is an instance of (38.1) and since we've take the closures of the instances of (38.1) as axioms, $\forall x(Px \to (Qx \to Px))$ is an axiom.

Witness to $\forall x P x \vdash \forall x(Qx \to Px)$

1.	$\forall x P x$	Premise
2.	$\forall x(Px \to (Qx \to Px))$	Closure of an instance of Axiom (38.1)
3.	$\forall x(Px \to (Qx \to Px)) \to$ $(\forall x P x \to \forall x(Qx \to Px))$	Instance of Axiom (39.3)
4.	$\forall x P x \to \forall x(Qx \to Px)$	from 2,3, by MP
5.	$\forall x(Qx \to Px)$	from 1,4, by MP

This sequence is a *bona fide* derivation of $\forall x(Qx \to Px)$ from $\forall x P x$, in the style of Frege and Hilbert. In this particular example, the *bona fide* derivation is actually shorter by three steps than the meta-derivation that appeals to GEN. Most of the time, however, the meta-derivations that invoke GEN are shorter than *bona fide* derivations that don't. Of course, the reasoning with GEN already looks a bit more straightforward than the reasoning without it. In any case, the two sequences described above show (a) how to use the metarule GEN to derive a universal claim, and (b) how to eliminate the use of GEN so as to derive the universal claim without it.

In light of the above facts, we shall take the liberty of reasoning with GEN as if it were a rule of inference instead of a metarule. The following example, in which we establish $\forall x P x \vdash \forall x(Qx \to Px)$, will be typical of the reasoning we use in the Appendix:

> From the premise $\forall x P x$ and the instance $\forall x P x \to (x\downarrow \to Px)$ of axiom (39.1), it follows that $x\downarrow \to Px$, by MP. From this result and the instance $x\downarrow$ of axiom (39.2), it follows that Px, by MP. From this last conclusion and the instance $Px \to (Qx \to Px)$ of axiom (38.1), it follows by MP that $Qx \to Px$. Since x isn't free in our premise, it follows that $\forall x(Qx \to Px)$, by GEN. ⋈

The above discussion should have made it clear just what has and has not been accomplished in this piece of reasoning. We sometimes deploy other metarules in just this way.

(**67**) **Remark:** Conventions Regarding Metarules. Although GEN was formulated to apply to $\vdash$, it also applies to $\vdash_\Box$. As noted earlier, the following version can be proved (exercise) by a trivial reworking of the justification for (66), since none of the reasoning involved an appeal to a modally fragile axiom:

- If $\Gamma \vdash_\Box \varphi$ and α doesn't occur free in any formula in Γ, then $\Gamma \vdash_\Box \forall \alpha \varphi$.

However, in what follows, we usually refrain from formulating metarules twice, with one form for $\vdash$ and a second form for $\vdash_\Box$. Instead, we adopt the conventions:

(.1) Whenever a metarule of inference is formulated generally, so as to apply to $\vdash$, we usually omit the statement of the metarule for the case of $\vdash_\Box$.

The only exception to this occurs when one of the conditions in the metarule specifically requires $\vdash_\Box$ instead of $\vdash$.

(.2) No metarule is to be adopted if the justification of the rule depends on a modally fragile axiom such as (43)★.

Though these conventions may be discussed on other occasions below, the following brief remarks may be sufficient for now. As noted previously, the justifications of metarules provided in the Appendix show how to convert reasoning with the metarules into *bona fide* derivations that don't use them. As long as the justification doesn't depend on a modally fragile axiom such as (43)★, a justification of a rule stated for $\vdash$ can be repurposed, with just a few obvious and trivial changes, to a justification of the analogous rule for $\vdash_\Box$. Thus, any any metarule of inference that applies to derivations and proofs *generally* (i.e., none of its conditions specifically involve $\vdash_\Box$) will be a metarule of inference that also applies to modally-strict derivations and proofs.

The next rule we consider, RN, contrasts with GEN because it is *not* a metarule that applies generally to all derivations and proofs. The antecedent of the Rule of Necessitation requires the existence of a modally-strict derivation or proof for the metarule to be applied.

(**68**) **Metarules:** Rule of Necessitation. The Rule of Necessitation (RN) is formulated in a way that prevents us from inferring the necessitation of a formula whose derivation or proof depends upon a modally fragile axiom. RN may be stated as follows:

- If there is a modally strict derivation of φ from zero or more premises, then there is a modally strict derivation of $\Box\varphi$ from the necessitations of the premises.

To formulate this statement precisely, we first introduce a metadefinition:

- For any set of formulas Γ, we use the notation $\Box\Gamma$ to refer to $\{\Box\psi \mid \psi \in \Gamma\}$.

So $\Box\Gamma$ is the set of formulas that results when a $\Box$ is prefixed to every formula in Γ. Then RN may be stated as follows:

Rule of Necessitation (**RN**)
If $\Gamma \vdash_\Box \varphi$, then $\Box\Gamma \vdash_\Box \Box\varphi$.

Rule of Necessitation (Weaker Form)[132]
If $\Gamma \vdash_\Box \varphi$, then $\Box\Gamma \vdash \Box\varphi$.

When $\Gamma = \varnothing$, RN reduces to:

[132] To see that the weaker form holds given the stronger form, assume $\Gamma \vdash_\Box \varphi$. Then by RN, $\Box\Gamma \vdash_\Box \Box\varphi$. But then by (62.1), it follows that $\Box\Gamma \vdash \Box\varphi$.

- If $\vdash_\Box \varphi$, then $\vdash_\Box \Box\varphi$.
- If $\vdash_\Box \varphi$, then $\vdash \Box\varphi$. (Weaker Form)

When we use RN to prove theorems, the reasoning almost always suggests that we are using the weaker form (since we rarely make it explicit that the conclusion is being derived by modally strict means). But, clearly, in what follows, we may use either form of RN *with the understanding* that any conclusions drawn via the metarule within a larger reasoning context do not affect the modal strictness of that context (or lack thereof). However, in the following examples, we use the stronger form, for purposes of illustration.

Here is an example of how Rule RN is to be applied, in which we derive φ ($= Qb$) from Γ ($= \{Pa, Pa \to Qb\}$) and then conclude by RN that $\Box\varphi$ is derivable, by modally strict means, from $\Box\Gamma$:

Example 1

1.	Pa	Premise
2.	$Pa \to Qb$	Premise
3.	Qb	from 1,2, by MP
4.	$Pa, Pa \to Qb \vdash_\Box Qb$	from 1–3, by df $\Gamma \vdash_\Box \varphi$ (60)
5.	$\Box Pa, \Box(Pa \to Qb) \vdash_\Box \Box Qb$	from 4, by RN

Lines 1–3 in this example constitute a witness to line 4 since (a) Qb follows by MP from two previous members of the sequence, both of which are in Γ, and (b) the derivation of Qb from the premises doesn't depend on a modally fragile axiom. RN then states that line 5 follows from line 4, so that we end up establishing that there is a modally strict derivation of $\Box Qb$ from $\Box Pa$ and $\Box(Pa \to Qb)$. Note that the conclusion on line 5 doesn't require that the premises Pa and $Pa \to Qb$ be true, nor that they be necessary. Even if the premises in a derivation are all contingently false (i.e., false but possibly true) and the conclusion is contingently false, RN will still assert that *if* there is a modally strict derivation of the conclusion from the premises, *then* there is a modally strict derivation of the necessitation of the conclusion from the necessitations of the premises. This in no way implies that the necessitations of the premises or conclusion are true.

Now, just as with GEN, the justification of RN in the Appendix shows us how to turn reasoning that appeals to RN into reasoning that does not. Although the sequence of formulas 1–4 in Example 1 isn't a witness to the derivability claim on line 5, the justification of RN in the Appendix shows us how to convert Example 1 into the following 5-element annotated sequence that is such a witness:

Witness to $\Box Pa, \Box(Pa \to Qb) \vdash_\Box \Box Qb$

1.	$\Box Pa$	Premise in $\Box\Gamma$
2.	$\Box(Pa \to Qb)$	Premise in $\Box\Gamma$
3.	$\Box(Pa \to Qb) \to (\Box Pa \to \Box Qb)$	Instance of Axiom (45.1)
4.	$\Box Pa \to \Box Qb$	from 2,3, by MP
5.	$\Box Qb$	from 1,4, by MP

This conversion works generally for any formulas φ and ψ: since there is a modally strict derivation of ψ from φ and $\varphi \to \psi$, there is a modally strict derivation of $\Box\psi$ from $\Box\varphi$ and $\Box(\varphi \to \psi)$.

Given the above discussion, it should be straightforward to see why we shall adopt the following, less formal style of reasoning in the Appendix when presented with a case like Example 1:

> Let $\Box Pa$ and $\Box(Pa \to Qb)$ be premises. Note that from the non-modal premises Pa and $Pa \to Qb$ it follows that Qb, by MP. Since this is a modally-strict derivation of Qb from Pa and $Pa \to Qb$, it follows (by modally strict reasoning) from our first two premises that $\Box Qb$, by RN. ⋈

In effect, we have reasoned by producing a modally strict 'sub-derivation' showing $Pa, Pa \to Qb \vdash_\Box Qb$, within the larger, modally strict derivation of $\Box Qb$ from $\Box Pa$ and $\Box(Pa \to Qb)$.

We now consider an example of RN involving quantifiers – one that involves a slight variant of the example we used to illustrate GEN. As an instance of RN we know: if there is a modally strict derivation of $\forall x(Qx \to Px)$ from $\forall xPx$, then there is a modally strict derivation of $\Box\forall x(Qx \to Px)$ from $\Box\forall xPx$:

Example 2

1.	$\forall xPx$	Premise
2.	$\forall xPx \to (x{\downarrow} \to Px)$	Instance of Axiom (39.1)
3.	$x{\downarrow} \to Px$	from 1,2, by MP
4.	$x{\downarrow}$	Instance of Axiom (39.2)
5.	Px	from 3,4, by MP
6.	$Px \to (Qx \to Px)$	Instance of Axiom (38.1)
7.	$Qx \to Px$	from 5,6, by MP
8.	$\forall xPx \vdash_\Box Qx \to Px$	from 1–7, by df $\Gamma \vdash_\Box \varphi$ (60)
9.	$\forall xPx \vdash_\Box \forall x(Qx \to Px)$	from 8, by GEN
10.	$\Box\forall xPx \vdash_\Box \Box\forall x(Qx \to Px)$	from 9, by RN

Here, lines 1–7 constitute a witness to line 8, which asserts that there is modally-strict derivation of $Qx \to Px$ from $\forall xPx$. On line 9, we apply the version of GEN that governs $\vdash_\Box$, which was discussed in (67). So line 9 satisfies the condition for the application of RN, which then implies the conclusion on line 10. The justification of RN itself shows us how to convert Example 2 into a witness for $\Box\forall xPx \vdash_\Box \Box\forall x(Qx \to Px)$:

Witness to $\Box\forall xPx \vdash_{\Box} \Box\forall x(Qx \to Px)$

1.	$\Box\forall xPx$	Premise
2.	$\Box[\forall x(Px \to (Qx \to Px)) \to$ $(\forall xPx \to \forall x(Qx \to Px))]$	Closure of an instance of Axiom (39.3)
3.	$\Box[\forall x(Px \to (Qx \to Px)) \to$ $(\forall xPx \to \forall x(Qx \to Px))] \to$ $(\Box\forall x(Px \to (Qx \to Px)) \to$ $\Box(\forall xPx \to \forall x(Qx \to Px)))$	Instance of Axiom (45.1)
4.	$\Box\forall x(Px \to (Qx \to Px)) \to$ $\Box(\forall xPx \to \forall x(Qx \to Px))$	from 2,3, by MP
5.	$\Box\forall x(Px \to (Qx \to Px))$	Closure of an instance of Axiom (38.1)
6.	$\Box(\forall xPx \to \forall x(Qx \to Px))$	from 4,5, by MP
7.	$\Box(\forall xPx \to \forall x(Qx \to Px)) \to$ $(\Box\forall xPx \to \Box\forall x(Qx \to Px))$	Instance of Axiom (45.1)
8.	$\Box\forall xPx \to \Box\forall x(Qx \to Px)$	from 6,7, by MP
9.	$\Box\forall x(Qx \to Px)$	from 1,8, by MP

This is a *bona fide* witness to the derivability claim since every line is either a necessary axiom, a premise, or follows from previous lines by MP. It should now be clear how the reasoning using GEN and RN in Example 2 is far easier to develop, or even grasp, when compared to the above reasoning. Indeed, we may compress the reasoning in Example 2 even further. The reasoning used in the Appendix for examples like this goes as follows:

> Assume $\Box\forall xPx$ as a global premise. From the local premise $\forall xPx$ and the instance $\forall xPx \to (x{\downarrow} \to Px)$ of axiom (39.1), it follows that $x{\downarrow} \to Px$, by MP. From this last result and the instance $x{\downarrow}$ of axiom (39.2), it follows that Px, by MP. From this last conclusion and the instance $Px \to (Qx \to Px)$ of axiom (38.1), it follows by MP that $Qx \to Px$. Since x isn't free in our local premise, it follows that $\forall x(Qx \to Px)$, by GEN. Since this constitutes a modally-strict derivation of $\forall x(Qx \to Px)$ from the local premise $\forall xPx$, it follows from our global premise $\Box\forall xPx$ that $\Box\forall x(Qx \to Px)$ (by modally strict means). $\bowtie$

Given the preceding discussion, this reasoning should be transparent; though it is not an actual derivation, it shows us how to construct one. The reasoning is a metaderivation that establishes the existence of a derivation. Consequently, we have a way to confirm derivability claims without producing actual derivations.

(**69**) **Remark:** Preserving GEN and RN When Extending The System. It is important to note that, should one wish to extend our system with new axioms, GEN and RN can easily be preserved as justified metarules. GEN remains valid as long as we axiomatically assert the universal closures of any new axioms we assert.[133] RN remains valid as long as we axiomatically assert the modal closures of any new necessary axioms we assert.[134]

(**70**) **Remark:** Modally Strict Reasoning, Modally Fragile Axioms, and Contingent Premises. Using an intuitive notion of truth (or if you prefer, a notion of truth relative to some fixed interpretation of our language), then we may say that a contingent formula is one that is neither necessarily true nor necessarily false (i.e., one that is both possibly true and possibly false). It is important to understand the conditions under which RN can be applied when reasoning from contingent premises. In particular, it is important to register the difference between:

(a) a derivation in which one of the premises, if there are any, is a contingent formula, but the derivation doesn't depend on any modally fragile axiom, and

(b) a derivation that depends upon a modally fragile *axiom*.

The first kind of derivation is *modally strict*, while the second kind is not.

For a scenario of type (a), start with Example 1 in the discussion of (68):

Example 1

1.	Pa	Premise
2.	$Pa \to Qb$	Premise
3.	Qb	from 1,2, by MP
4.	$Pa, Pa \to Qb \vdash_{\Box} Qb$	from 1–3, by df $\Gamma \vdash_{\Box} \varphi$ (60)
5.	$\Box Pa, \Box(Pa \to Qb) \vdash_{\Box} \Box Qb$	from 4, by RN

In Example 1, the conclusion on line 4, i.e., that there exists a modally strict derivation of Qb from Pa and $Pa \to Qb$, holds *even if* one or more of Pa, $Pa \to Qb$, and Qb are contingent. The modal strictness of a derivation from

[133] The justification of GEN in the Appendix is by induction on the length of a derivation of $\Gamma \vdash \varphi$. The Base Case considers the two ways in which such a derivation could consist of a single formula, namely, either φ is an axiom or φ is in Γ. When φ is an axiom, then since we've taken all the universal closures of axioms as axioms, it follows that $\forall\alpha\varphi$ is an axiom as well. So $\forall\alpha\varphi$ is a theorem (63.1) and so is derivable from Γ (63.3).

If we extend our system with new axioms, the reasoning in this case will be preserved as long as we always take their universal closures as axioms as well.

[134] Just as with GEN, the justification of RN in the Appendix is by induction on the length of a derivation, though in this case, on the length of a derivation $\Gamma \vdash_{\Box} \varphi$. The Base Case considers the two ways in which such a derivation could consist of a single formula, namely, either φ is an necessary axiom or φ is in Γ. When φ is necessary axiom, then since we've taken all the modal closures of axioms as axioms, it follows that $\Box\varphi$ is an axiom as well. So $\Box\varphi$ is a modally strict theorem (63.1) and becomes derivable from Γ by modally strict means (63.3).

If we extend our system with new necessary axioms, the reasoning in this case will be preserved as long as we always take their modal closures as axioms as well.

contingent premises to a contingent conclusion is *not* undermined as long as no modally fragile axioms (or, more generally, no $\star$-theorems) are used in the derivation. So RN can be applied on line 5, for it requires only that there be a modally strict derivation of Qb from Pa and $Pa \to Qb$.

This applies even for *conditional proof*, which is introduced later in (75). Suppose we reason by conditional proof to establish, as a theorem, $Pa \to (Pa \vee Qb)$. That is, suppose we *assume* Pa, derive $Pa \vee Qb$, and then conclude $Pa \to (Pa \vee Qb)$ by conditional proof. Then even if Pa is a contingent claim, it is *discharged* when we apply conditional proof and so the theorem no longer depends on a contingency; the theorem in question is modally strict. For a more interesting example, see the discussion in (218).

Contrast this with a scenario of type (b). Suppose that we have *extended our system* with the modally fragile axiom Pa and annotated it with a $\star$, say because (we know) it is not necessarily true. (We can even suppose we've added $\neg\Box Pa$ as a further axiom to make it clear that Pa is contingent.) By adding Pa as a new, modally-fragile axiom, we have extended (the definitions of) our deductive system and consequently, of the derivability conditions $\vdash$ and $\vdash_\Box$. As such, we have a new deductive relationship between $Pa \to Qb$ and Qb. Qb becomes derivable from the sole premise $Pa \to Qb$ but not by modally strict means, i.e., $Pa \to Qb \vdash Qb$ but not $Pa \to Qb \vdash_\Box Qb$. A derivation that bears witness to $Pa \to Qb \vdash Qb$ will fail to be modally strict because the modally fragile axiom Pa is required in the derivation. So we can't apply RN to conclude $\Box(Pa \to Qb) \vdash \Box Qb$.[135] To take another example, suppose that we have taken $\neg\Box Pa$ as a necessary axiom along with the modally fragile axiom Pa. Then, from the $\star$-axiom Pa and the necessary axiom $\neg\Box Pa$, we may derive, by definition (213.1), that *ContingentlyTrue*(Pa). This latter becomes a $\star$-theorem since its proof depends on a modally fragile axiom.

In what follows, it will be essential to distinguish (a) modally strict derivations from contingent premises and (b) reasoning that is not modally strict because it depends upon a modally fragile axiom.

(**71**) **Remark:** Digression on the Converse of Weak RN. For the purposes of this remark, we consider only the form of Rule RN when the premise set Γ is empty. When we introduced RN in (68), we noted that when the premise set is empty, the weak form of RN asserts:

- If $\vdash_\Box \varphi$, then $\vdash \Box\varphi$.

We indicated that, for simplicity, we shall typically cite only this weaker form with the understanding that any modal strictness of the reasoning context is preserved.

[135] Note, however, that the deductive relationship $Pa, Pa \to Qb \vdash_\Box Qb$ *still holds* and we can still apply RN to conclude $\Box Pa, \Box(Pa \to Qb) \vdash \Box Qb$. But we may not have much use for this fact if we've added $\neg\Box Pa$ as as axiom.

However, it is important to distinguish the *converse* of RN from the *converse* of the weaker form. While the converse of RN is guaranteed to hold, the converse of the weaker form is not. The converse of RN is: if $\vdash_\Box \Box\varphi$, then $\vdash_\Box \varphi$. Clearly, this holds. For assume $\vdash_\Box \Box\varphi$. Since instances of the T schema $\Box\varphi \to \varphi$ are necessary axioms (45.2), it follows by (63.1) that $\vdash_\Box \Box\varphi \to \varphi$. But from this and our assumption that $\vdash_\Box \Box\varphi$, it follows by (63.5) that $\vdash_\Box \varphi$.

By contrast, the converse of the weaker form of RN is not valid or, rather, not *robustly* valid under reasonable extensions of the system. The converse asserts: if $\vdash \Box\varphi$, then $\vdash_\Box \varphi$. To see a counterexample to this claim, consider a very natural extension of our system (or, if you prefer, a model in which certain new axioms hold). Consider the scenario in the previous Remark, in which we extended our system with the contingent axiom Pa ('object a exemplifies property P') and where the contingency is a direct consequence of an additional axiom, namely, that $\Diamond\neg Pa$. So Pa would be a modally fragile axiom and marked with a $\star$. Then our theory would guarantee, by theorem (250), that there is a unique abstract object that encodes all and only the properties that a exemplifies. So we would be able to show that this abstract object, call it $\boldsymbol{c}_a$, encodes P, i.e., that $\boldsymbol{c}_a P$. This conclusion would be a $\star$-theorem since it was derived with the help of a modally fragile axiom. But note that by the axiom for the rigidity of encoding (51), it would also follow that $\boldsymbol{c}_a$ necessarily encodes P. So $\Box\boldsymbol{c}_a P$ becomes a theorem and, indeed, a $\star$-theorem. Thus, we would have $\vdash \Box\boldsymbol{c}_a P$. But there would be no modally strict proof of $\boldsymbol{c}_a P$; the proof of $\boldsymbol{c}_a P$ requires an appeal to the modally fragile axiom that Pa. Hence, we would have a formula φ, namely $\boldsymbol{c}_a P$, such that $\vdash \Box\varphi$ but not $\vdash_\Box \varphi$. Thus, the converse of the weaker form of RN would fail. An example of this kind is discussed in some detail in (257).

However, if we fix the axioms in terms of the set Λ as defined in (59), which includes only the axioms presented in Chapter 8, then one can show that the converse of weak RN is valid. This was pointed out and proved independently by Daniel Kirchner and Daniel West. However, we postpone further discussion here, since the proof of the converse of weak RN becomes extremely easy once we establish a Fact that is discussed and proved in Remark (137).

9.4 The Inferential Role of Definitions

Definitions play an important part in the proof of many theorems; few theorems are expressed using only primitive notions or rest solely on axioms that are expressed using only primitive notions. However, the classical theory of definitions has to be modified in a number of ways if we are to obtain a serviceable theory of definitions for a system like the present one, which has far greater expressive power than the systems for which the classical theory was developed. We've already discussed a number of issues about the formulation and understanding of definitions, in Remarks (17), (27), (28), (31), (32), and (36). We also developed conventions in Remark (17), governing definitions. In particular, Convention (17.2) allows us to use object-language variables in definitions with the understanding that they function as metavariables.

In this section, however, we describe how definitions impact our reasoning system. Some readers may find it useful to skip ahead to Remark (282), which offers one important reason why we don't use the classical theory of definitions, in which new individual terms are introduced using *definitions by equivalence*. In what follows, we rigorously distinguish the inferential roles of definitions-by-≡ and definitions-by-=. The resulting theory avoids the problems noted not only in Remark (282), but in Remark (283) as well. By the end of this section, the reader should have a good grasp on what is meant when we say 'So, by definition, ... ' during the course of reasoning in proofs and derivations.

(72) **Primitive Metarule:** The Inferential Role of Definitions by Equivalence. Though the general case of a definition-by-≡ has the form $\varphi(\alpha_1,\ldots,\alpha_n) \equiv_{df} \psi(\alpha_1,\ldots,\alpha_n)$, let us abbreviate this more simply as $\varphi \equiv_{df} \psi$ and suppose that this represents any valid instance of the definition, i.e., any instance having the form $\varphi(\tau_1,\ldots,\tau_n) \equiv_{df} \psi(\tau_1,\ldots,\tau_n)$, where $\tau_1,\ldots,\tau_n$ are substitutable for $\alpha_1,\ldots,\alpha_n$, respectively, in ψ.

Note that since the equivalence symbol (≡) is defined using a definition-by-equivalence (18.3), we cannot present the inferential role of definitions-by-equivalence in terms of inferences involving the biconditional. For this wouldn't be useful at this point; as yet, we haven't yet proved the tautologies (governing ≡ and &) that allow us to reason from biconditionals. So, in the first instance, we specify the inferential role in terms of primitives of the language, i.e., in terms of conditionals.

We therefore stipulate that the inferential role of a definition-by-≡ is captured by the following primitive metarule of inference:

> **Rule of Definition by Equivalence**
> A definition-by-≡ of the form $\varphi \equiv_{df} \psi$ introduces the closures of $\varphi \to \psi$ and $\psi \to \varphi$ as necessary axioms.

So given the definition $\varphi \equiv_{df} \psi$, the closures of $\varphi \to \psi$ and $\psi \to \varphi$ (a) become modally strict theorems, by (63.1), and (b) can be derived (by modally strict means) from any set of premises Γ, by (63.3). Thus, the rule can be applied in any reasoning context without affecting the modal strictness of the reasoning.

Once we have established the tautologies governing & and ≡, we will then be in a position to establish that $\varphi \equiv_{df} \psi$ implies $\varphi \equiv \psi$ as a modally strict theorem; see (90.1). Moreover, in the discussion immediately following the presentation of (90.1), we explain how the definition $\varphi \equiv_{df} \psi$ wil also yield the *closures* of $\varphi \equiv \psi$ as modally strict theorems.

(**73**) **Primitive Metarule:** The Inferential Role of Definitions by Identity. Those familiar with the classical theory of definitions and the problem of non-significant terms will recognize that the inferential role of definitions-by-= has to be carefully formulated. A hint as to why this is so was offered at the end of Remark (27), but those not familiar with the issues here may find Remarks (282) and (283) useful. The latter Remark, in particular, takes the reader step-by-step through the motivation and justification of the following metarule as it applies to definitions with one free variable in the definiens and definiendum.

Let $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$ be a definition-by-= in which the variables $\alpha_1,\ldots,\alpha_n$ occur free ($n \geq 0$). So when $n = 0$ and no variables occur free in σ, the definition has the form $\tau =_{df} \sigma$ and introduces the new constant τ. Now let us use $\sigma(\tau_1,\ldots,\tau_n)$ and $\tau(\tau_1,\ldots,\tau_n)$ to abbreviate $\sigma^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$ and $\tau^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$, respectively, so that $\sigma(\tau_1,\ldots,\tau_n)$ and $\tau(\tau_1,\ldots,\tau_n)$ are the result of uniformly substituting τ_i for the free occurrences of α_i in σ and τ respectively ($1 \leq i \leq n$). Then we may state the inferential role of a definition-by-= as follows:

> **Rule of Definition by Identity**
> Whenever $\tau_1,\ldots,\tau_n$ are substitutable for $\alpha_1,\ldots,\alpha_n$, respectively, in σ, then a definition-by-= of the form $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$ introduces the closures of the following, necessary axiom schema:
>
> $$(\omega)\ \ (\sigma(\tau_1,\ldots,\tau_n)\downarrow \rightarrow \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n))\ \&$$
> $$(\neg\sigma(\tau_1,\ldots,\tau_n)\downarrow \rightarrow \neg\tau(\tau_1,\ldots,\tau_n)\downarrow)$$

Ignoring closures, the rule says that the inferential role of an *instance* of a definition-by-= is to introduce a necessary axiom that asserts: (a) if the definiens is significant, then an identity holds between the definiendum and definiens, and (b) if the definiens is empty, then the definiendum is empty.[136] A full discussion motivating and justifying this rule is given in (283), though see also (282).

As we shall see, the Rule of Definition by Identity will yield, in (120.2), the classical introduction and elimination rules for the definiendum when the definiens is significant. Moreover, the Rule of Definition by Identity finesses the problem of 'conditional definitions', such as division by zero in real number theory. But a full discussion of this and other interesting issues that this rule gives rise to will be reserved for Remark (283), where we fully motivate

[136] Stated a bit more strictly, the inferential role of an *instance* of a definition-by-= is to introduce a necessary axiom that asserts: (a) if the definiens, when applied to suitable arguments $\tau_1,\ldots,\tau_n$, is significant, then an identity holds between the definiendum and definiens when both are applied to those arguments, and (b) if the definiens is empty, when applied to the arguments $\tau_1,\ldots,\tau_n$, then the definiendum, when applied to those arguments, is empty.

the rule as stated above.[137]

When $n = 0$, the definiens in a definition-by-= is a closed term and the definiendum is a new constant. So our rule asserts that the definition $\tau =_{df} \sigma$ introduces the necessary axiom:

$$(\sigma\downarrow \rightarrow \tau = \sigma) \;\&\; (\neg\sigma\downarrow \rightarrow \neg\tau\downarrow)$$

Once we have rules in place for reasoning from conjunctions, we'll see both that the above immediately yields $\tau = \sigma$ as a $\Box$-theorem whenever we can establish $\sigma\downarrow$ as a $\Box$-theorem, and yields $\tau = \sigma$ as a $\star$-theorem whenever we can establish $\sigma\downarrow$ as a $\star$-theorem. In both cases, τ and σ become substitutable for one another in any context, though in the latter case, the substitution will undermine the modal strictness of the reasoning.

9.5 The Theory of Negations and Conditionals

(**74**) **Theorems:** A Useful Fact. The following fact is derivable and is crucial to the proof of the Deduction Theorem:

$$\varphi \rightarrow \varphi$$

Although the notion of a *tautology* is a semantic notion and isn't officially defined in our formal system, we saw in Section 6.2 that the notion can be precisely defined if one takes on board the required semantic notions. It won't hurt, therefore, if we use the notion unofficially and label the above claim a tautology. Other tautologies will be derived below. As we will see, all tautologies are derivable, but it will be some time before we have assembled all the facts needed to prove this metatheoretic fact.

(**75**) **Metarule:** Deduction Theorem and Conditional Proof (CP). If there is a derivation of ψ from a set of premises Γ together with an additional premise φ, then there is a derivation of $\varphi \rightarrow \psi$ from Γ:

Rule CP
If $\Gamma, \varphi \vdash \psi$, then $\Gamma \vdash (\varphi \rightarrow \psi)$.

[137] To anticipate the issue of conditional definition, consider that in real number theory, where x, y, z range over real numbers, one might define the operation of division, i.e., x/y, as follows: if it is provable or assumed that $y \neq 0$, then $x/y =_{df} \imath z(x = y \cdot z)$. Without the antecedent condition "if it is provable or assumed that $y \neq 0$", the description $\imath z(x = y \cdot z)$ can fail to denote: there is no such z when x is any positive number and $y = 0$, and there is no *unique* such z when $x = 0$ and $y = 0$. But given our Rule of Definition by Identity, the definition yields axioms of the following form:

$$(\imath z(x = y \cdot z)\downarrow \rightarrow (x/y = \imath z(x = y \cdot z))) \;\&\; (\neg \imath z(x = y \cdot z)\downarrow) \rightarrow \neg(x/y)\downarrow)$$

So when $y = 0$, one can establish, for any x, that $\neg \imath z(x = 0 \cdot z)\downarrow$, and thereby conclude $\neg(x/0)\downarrow$. One doesn't have to use 'conditional definitions' that conditionally extend the language only when some condition is provable.

By convention (67.1), we omit the formulation of the analogous rule for $\vdash_{\Box}$. And hereafter, we shall not remark upon such omissions unless a special case calls for some such remark. Rule CP is most-often used when $\Gamma = \varnothing$:

If $\varphi \vdash \psi$, then $\vdash \varphi \to \psi$.

When we cite this metarule in the proof of other metarules, we reference it as the *Deduction Theorem*. However, we shall adopt the following convention: during the course of reasoning, once we have produced a derivation of ψ from φ, then instead of concluding $\vdash \varphi \to \psi$ and citing the Deduction Theorem, we shall *infer* $\varphi \to \psi$ and cite *Conditional Proof* (CP). The proof of the Deduction Theorem in the Appendix guarantees that we can indeed construct a proof of the conditional $\varphi \to \psi$ once we have derived ψ from φ.[138]

(**76**) **Metarules:** Corollaries to the Deduction Theorem. The following metarules are immediate consequences of the Deduction Theorem. They help us to prove the tautologies in (77), (85), and (88). Recall that 'Γ_1, Γ_2' indicates '$\Gamma_1 \cup \Gamma_2$':

(.1) If $\Gamma_1 \vdash \varphi \to \psi$ and $\Gamma_2 \vdash \psi \to \chi$, then $\Gamma_1, \Gamma_2 \vdash \varphi \to \chi$

(.2) If $\Gamma_1 \vdash \varphi \to (\psi \to \chi)$ and $\Gamma_2 \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \varphi \to \chi$

It is interesting that the above metarules have the following Variant forms, respectively:

(.3) $\varphi \to \psi,\ \psi \to \chi \vdash \varphi \to \chi$ [**Hypothetical Syllogism**]

(.4) $\varphi \to (\psi \to \chi),\ \psi \vdash \varphi \to \chi$

(.3) is a Variant of (.1) because we can derive each from the other.[139] Similarly, (.4) is a Variant of (.2).

---

[138] Metatheorem ⟨6.7⟩, which is proved in the Appendix to Chapter 6, establishes that $\Gamma, \varphi \models \psi$ if and only if $\Gamma \models (\varphi \to \psi)$. Furthermore, Metatheorem ⟨6.8⟩, which is also proved in the Appendix to Chapter 6, establishes that $\varphi \models \psi$ if and only if $\models \varphi \to \psi$.

[139] Here is a proof. ($\hookrightarrow$) Assume (.1), i.e., if $\Gamma_1 \vdash \varphi \to \psi$ and $\Gamma_2 \vdash \psi \to \chi$, then $\Gamma_1, \Gamma_2 \vdash \varphi \to \chi$. We want to show $\varphi \to \psi,\ \psi \to \chi \vdash \varphi \to \chi$. If we let Γ_1 be $\{\varphi \to \psi\}$, then since by (63.2) we know $\varphi \to \psi \vdash \varphi \to \psi$, we have $\Gamma_1 \vdash \varphi \to \psi$. By similar reasoning, if we let Γ_2 be $\{\psi \to \chi\}$, then we have $\Gamma_2 \vdash \psi \to \chi$. Hence, by (.1), it follows that $\Gamma_1, \Gamma_2 \vdash \varphi \to \chi$. But, this is just $\varphi \to \psi, \psi \to \chi \vdash \varphi \to \chi$. ($\hookleftarrow$) Assume (.3), i.e., $\varphi \to \psi,\ \psi \to \chi \vdash \varphi \to \chi$. Then by (63.7), it follows that:

$\Gamma_1, \Gamma_2, \varphi \to \psi,\ \psi \to \chi \vdash \varphi \to \chi$

From this, by two applications of the Deduction Theorem, we have:

(ϑ) $\Gamma_1, \Gamma_2 \vdash (\varphi \to \psi) \to ((\psi \to \chi) \to (\varphi \to \chi))$

Now to show (.1), assume $\Gamma_1 \vdash \varphi \to \psi$ and $\Gamma_2 \vdash \psi \to \chi$. So by (63.7), it follows, respectively, that:

(a) $\Gamma_1, \Gamma_2 \vdash \varphi \to \psi$

(b) $\Gamma_1, \Gamma_2 \vdash \psi \to \chi$

By (a) and (ϑ), it follows by (63.5) that:

(ξ) $\Gamma_1, \Gamma_2 \vdash (\psi \to \chi) \to (\varphi \to \chi)$

From (ξ) and (b), it follows by (63.5) that $\Gamma_1, \Gamma_2 \vdash \varphi \to \chi$. ⋈

Note that the Variants (.3) and (.4) are somewhat different from the stated metarules (.1) and (.2): the Variants don't have the form of a conditional but instead simply assert the existence of a derivation. Of course they can be put into the traditional metarule form by conditionalizing them upon the triviality "If any condition holds" or "Under all conditions". But, given that these metarules hold without preconditions, we may consider them *derived rules of inference*, i.e., rules of inference, like Modus Ponens, that allow us to infer formulas from formulas. Thus, (.3), for example, can be reconceived as a derived rule and not just a metarule. We may justifiably use this rule within derivations and consider the result to be a *bona fide* derivation. The justification of (.3) in the Appendix establishes that any derivation that yields a conclusion by an application of the above derived rule of Hypothetical Syllogism can be converted to a derivation in which this rule isn't used.

This pattern, of taking the unconditional variants of metarules to be derived rules, will occur often in what follows; many of the derived metarules for reasoning with negation and conditionals have unconditional variants that will be regarded as derived rules. In the Appendix, however, we typically reason with the variant, derived rule whenever it is available.

(**77**) **Theorems:** More Useful Tautologies. The tautologies listed below (and their proofs) follow the presentation in Mendelson 1964 [1997, 38–40, Lemma 1.11). We present them as a group because they are needed in the Appendix to this chapter to establish Lemma ⟨9.1⟩ and Metatheorem ⟨9.2⟩, i.e., that every tautology is derivable.

(.1) $\neg\neg\varphi \rightarrow \varphi$

(.2) $\varphi \rightarrow \neg\neg\varphi$

(.3) $\neg\varphi \rightarrow (\varphi \rightarrow \psi)$

(.4) $(\neg\psi \rightarrow \neg\varphi) \rightarrow (\varphi \rightarrow \psi)$

(.5) $(\varphi \rightarrow \psi) \rightarrow (\neg\psi \rightarrow \neg\varphi)$

(.6) $(\varphi \rightarrow \neg\psi) \rightarrow (\psi \rightarrow \neg\varphi)$

(.7) $(\neg\varphi \rightarrow \psi) \rightarrow (\neg\psi \rightarrow \varphi)$

(.8) $\varphi \rightarrow (\neg\psi \rightarrow \neg(\varphi \rightarrow \psi))$

(.9) $(\varphi \rightarrow \psi) \rightarrow ((\neg\varphi \rightarrow \psi) \rightarrow \psi)$

(.10) $(\varphi \rightarrow \neg\psi) \rightarrow ((\varphi \rightarrow \psi) \rightarrow \neg\varphi)$

(.1) and (.2) are used to derive Double Negation rules (78); (.5) is used to prove Modus Tollens (79); and (.10) is used to prove a form of reductio ad absurdum.

(78) **Metarules/Derived Rules:** Double Negation. It is easy to derive introduction and elimination metarules for double negation, along with their corresponding derived rules:

(.1) **Double Negation Introduction** (Rule $\neg\neg$I):

If $\Gamma \vdash \varphi$, then $\Gamma \vdash \neg\neg\varphi$ [Variant: $\varphi \vdash \neg\neg\varphi$]

(.2) **Double Negation Elimination** (Rule $\neg\neg$E):

If $\Gamma \vdash \neg\neg\varphi$, then $\Gamma \vdash \varphi$ [Variant: $\neg\neg\varphi \vdash \varphi$]

(79) **Metarules/Derived Rules:** Modus Tollens. We formulate Modus Tollens (MT) as two metarules:

Rules of Modus Tollens (**MT**)
(.1) If $\Gamma_1 \vdash (\varphi \rightarrow \psi)$ and $\Gamma_2 \vdash \neg\psi$, then $\Gamma_1, \Gamma_2 \vdash \neg\varphi$
[Variant: $\varphi \rightarrow \psi,\ \neg\psi\ \vdash\ \neg\varphi$]
(.2) If $\Gamma_1 \vdash (\varphi \rightarrow \neg\psi)$ and $\Gamma_2 \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \neg\varphi$
[Variant: $\varphi \rightarrow \neg\psi,\ \psi\ \vdash\ \neg\varphi$]

The Variants are equivalent, by reasoning analogous to that in footnote 139. In light of the discussion at the end of (76), the Variants may be conceived as the well-known derived rules.

(80) **Metarules/Derived Rules:** Contraposition. These metarules also come in two forms:

Rules of Contraposition
(.1) $\Gamma \vdash \varphi \rightarrow \psi$ if and only if $\Gamma \vdash \neg\psi \rightarrow \neg\varphi$
[Variant: $\varphi \rightarrow \psi \dashv\vdash \neg\psi \rightarrow \neg\varphi$]
(.2) $\Gamma \vdash \varphi \rightarrow \neg\psi$ if and only if $\Gamma \vdash \psi \rightarrow \neg\varphi$
[Variant: $\varphi \rightarrow \neg\psi \dashv\vdash \psi \rightarrow \neg\varphi$

In the Variants, $\chi \dashv\vdash \theta$ ('χ is interderivable with θ') is simply means $\chi \vdash \theta$ and $\theta \vdash \chi$.[140] Given the discussion at the end of (76), we typically use the derived rules in proofs.

[140] Although the reasoning is again analogous to that in footnote 139, we show here the left-to-right direction of (.1) is equivalent to the variant $\varphi \rightarrow \psi \vdash \neg\psi \rightarrow \neg\varphi$. ($\hookrightarrow$) Assume metarule (.1): if $\Gamma \vdash \varphi \rightarrow \psi$, then $\Gamma \vdash \neg\psi \rightarrow \neg\varphi$. Now let Γ be $\{\varphi \rightarrow \psi\}$. Then we have $\Gamma \vdash \varphi \rightarrow \psi$, by the special case of (63.2). But then it follows from our assumption that $\Gamma \vdash \neg\psi \rightarrow \neg\varphi$, i.e., $\varphi \rightarrow \psi \vdash \neg\psi \rightarrow \neg\varphi$. ($\hookleftarrow$) Assume $\varphi \rightarrow \psi \vdash \neg\psi \rightarrow \neg\varphi$. Then by (63.7), it follows that $\Gamma, \varphi \rightarrow \psi \vdash \neg\psi \rightarrow \neg\varphi$. By the Deduction Theorem, it follows that $\Gamma \vdash (\varphi \rightarrow \psi) \rightarrow (\neg\psi \rightarrow \neg\varphi)$. But from this fact, we can derive $\Gamma \vdash \neg\psi \rightarrow \neg\varphi$ from the assumption that $\Gamma \vdash \varphi \rightarrow \psi$, by (63.5).

We leave the other direction, and the proof of the equivalence of (.2) and its variant, as exercises.

(**81**) **Metarules/Derived Rules:** Reductio Ad Absurdum. Two classic forms of Reductio Ad Absurdum (RAA) are formulated as follows:

Rules of Reductio Ad Absurdum (**RAA**)
(.1) If $\Gamma_1, \neg\varphi \vdash \neg\psi$ and $\Gamma_2, \neg\varphi \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \varphi$
[Variant: $\neg\varphi \to \neg\psi, \neg\varphi \to \psi \vdash \varphi$]
(.2) If $\Gamma_1, \varphi \vdash \neg\psi$ and $\Gamma_2, \varphi \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \neg\varphi$
[Variant: $\varphi \to \neg\psi, \varphi \to \psi \vdash \neg\varphi$]

The variant versions are, respectively, provably equivalent to their more general formulations.[141] We may therefore use them as derived rules, if needed.

(**82**) **Remark:** Introduction and Elimination Rules for $\to$ and $\neg$. Note that the metarules for the introduction and elimination of $\to$ and $\neg$ have already been presented. (63.5) is the metarule for $\to$Elimination ($\to$E); the Deduction Theorem (75) and its corollaries (76) are metarules for $\to$Introduction ($\to$I). Reductio Ad Absurdum, when formulated as in (81.1), is a metarule for $\neg$Elimination ($\neg$E), and when formulated as in (81.2), is a metarule for $\neg$Introduction ($\neg$I). And we've formulated the introduction and elimination rules for double negation in (78).

We now work our ways towards the formulation of the introduction and elimination metarules for &, $\vee$, and $\equiv$. We also state their variant, derived rules when applicable, though we leave the proof that they are equivalent to the stated metarules for the reader. We also assume that, in each case, the $\vdash_\Box$ form of the metarules and derived rules can be easily proved using the $\vdash$ form as a guide.

(**83**) **Theorems:** Principle of Excluded Middle.

[141] We can show that (.1) is equivalent to the Variant as follows. ($\hookrightarrow$) Assume (.1): if $\Gamma_1, \neg\varphi \vdash \neg\psi$ and $\Gamma_2, \neg\varphi \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \varphi$. Now to derive the variant, note that if we let $\Gamma_1 = \{\neg\varphi \to \neg\psi\}$ and $\Gamma_2 = \{\neg\varphi \to \psi\}$, then we know by MP both that $\Gamma_1, \neg\varphi \vdash \neg\psi$ and $\Gamma_2, \neg\varphi \vdash \psi$. Hence by our assumption, $\Gamma_1, \Gamma_2 \vdash \varphi$. ($\hookleftarrow$) Assume the variant version, i.e., $\neg\varphi \to \neg\psi, \neg\varphi \to \psi \vdash \varphi$. By (63.7), it follows that:

$$\Gamma_1, \Gamma_2, \neg\varphi \to \neg\psi, \neg\varphi \to \psi \vdash \varphi$$

So by two applications of the Deduction Theorem, we know:

(ϑ) $\Gamma_1, \Gamma_2 \vdash (\neg\varphi \to \neg\psi) \to ((\neg\varphi \to \psi) \to \varphi)$

Now to show (81.1), assume $\Gamma_1, \neg\varphi \vdash \neg\psi$ and $\Gamma_2, \neg\varphi \vdash \psi$. Then by (63.7), it follows, respectively, that $\Gamma_1, \Gamma_2, \neg\varphi \vdash \neg\psi$ and $\Gamma_1, \Gamma_2, \neg\varphi \vdash \psi$. By applying the Deduction Theorem to each of these, we obtain, respectively:

(ξ) $\Gamma_1, \Gamma_2 \vdash \neg\varphi \to \neg\psi$

(ζ) $\Gamma_1, \Gamma_2 \vdash \neg\varphi \to \psi$

But from (ϑ) and (ξ), it follows by (63.5) that:

$$\Gamma_1, \Gamma_2 \vdash (\neg\varphi \to \psi) \to \varphi$$

And from this last conclusion and (ζ) it follows again by (63.5) that $\Gamma_1, \Gamma_2 \vdash \varphi$. $\bowtie$

We leave the proof of the equivalence of (.2) and its variant as an exercise.

$$\varphi \vee \neg\varphi$$

The proof in the Appendix of this classical principle involves an appeal to a previously established theorem (74) and our Rule of Definition by Equivalence (72).

(**84**) **Theorems:** Principle of Noncontradiction.

$$\neg(\varphi \,\&\, \neg\varphi)$$

The proof in the Appendix of (.1) goes by way of a reductio, and involves an appeal to a previously established theorem (77.2), the definition of & (18.1) and the Rule of Definition by Equivalence (72).

(**85**) **Theorems:** Basic Tautologies Governing Conjunction and Disjunction.

(.1) $(\varphi \,\&\, \psi) \to \varphi$ (Conjunction Simplification)

(.2) $(\varphi \,\&\, \psi) \to \psi$ (Conjunction Simplification)

(.3) $\varphi \to (\varphi \vee \psi)$ (Disjunction Addition)

(.4) $\psi \to (\varphi \vee \psi)$ (Disjunction Addition)

(.5) $\varphi \to (\psi \to (\varphi \,\&\, \psi))$ (Adjunction)

(.6) $(\varphi \,\&\, \varphi) \equiv \varphi$ (Idempotence of &)

(.7) $(\varphi \vee \varphi) \equiv \varphi$ (Idempotence of $\vee$)

(**86**) **Metarules/Derived Rules:** Introduction and Elimination Rules for Conjunction and Disjunction.

(.1) &Introduction (&I):

If $\Gamma_1 \vdash \varphi$ and $\Gamma_2 \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \varphi \,\&\, \psi$ [Variant: $\varphi, \psi \vdash \varphi \,\&\, \psi$]

(.2) &Elimination (&E):

(.a) If $\Gamma \vdash \varphi \,\&\psi$, then $\Gamma \vdash \varphi$ [Variant: $\varphi \,\&\, \psi \vdash \varphi$]

(.b) If $\Gamma \vdash \varphi \,\&\psi$, then $\Gamma \vdash \psi$ [Variant: $\varphi \,\&\, \psi \vdash \psi$]

(.3) $\vee$Introduction ($\vee$I):

(.a) If $\Gamma \vdash \varphi$, then $\Gamma \vdash \varphi \vee \psi$ [Variant: $\varphi \vdash \varphi \vee \psi$]

(.b) If $\Gamma \vdash \psi$, then $\Gamma \vdash \varphi \vee \psi$ [Variant: $\psi \vdash \varphi \vee \psi$]

(.c) Disjunctive Syllogism:
If $\Gamma_1 \vdash \varphi \vee \psi$, $\Gamma_2 \vdash \varphi \to \chi$, and $\Gamma_3 \vdash \psi \to \theta$, then $\Gamma_1, \Gamma_2, \Gamma_3 \vdash \chi \vee \theta$
[Variant: $\varphi \vee \psi$, $\varphi \to \chi$, $\psi \to \theta \vdash \chi \vee \theta$]

(.4) $\vee$Elimination ($\vee$E):

(.a) Reasoning by Cases:
If $\Gamma_1 \vdash \varphi \vee \psi$, $\Gamma_2 \vdash \varphi \rightarrow \chi$, and $\Gamma_3 \vdash \psi \rightarrow \chi$, then $\Gamma_1, \Gamma_2, \Gamma_3 \vdash \chi$
[Variant: $\varphi \vee \psi,\ \varphi \rightarrow \chi,\ \psi \rightarrow \chi \vdash \chi$]

(.b) Disjunctive Syllogism (alternative form):
If $\Gamma_1 \vdash \varphi \vee \psi$ and $\Gamma_2 \vdash \neg\varphi$, then $\Gamma_1, \Gamma_2 \vdash \psi$ [Variant: $\varphi \vee \psi,\ \neg\varphi \vdash \psi$]

(.c) Disjunctive Syllogism (alternative form):
If $\Gamma_1 \vdash \varphi \vee \psi$ and $\Gamma_2 \vdash \neg\psi$, then $\Gamma_1, \Gamma_2 \vdash \varphi$ [Variant: $\varphi \vee \psi,\ \neg\psi \vdash \varphi$]

(**87**) **Metarules/Derived Rules**. Classical and Alternative Forms of RAA. In systems where & is a primitive connective, classical Reductio Ad Absurdum is stated as a primitive metarule which asserts that if there is a derivation of a contradiction from some premises and an assumption, then there is a derivation of the negation of the assumption from the premises alone. But in our system, Reductio is derivable as a metarule, though only once the above theorems and rules are in place:

(.1) If $\Gamma, \neg\varphi \vdash \psi \,\&\, \neg\psi$, then $\Gamma \vdash \varphi$

(.2) If $\Gamma, \varphi \vdash \psi \,\&\, \neg\psi$, then $\Gamma \vdash \neg\varphi$

The reader should also confirm that Reductio Ad Absurdum also has the following forms:

(.3) If $\Gamma,\ \varphi,\ \neg\psi \vdash \neg\varphi$, then $\Gamma, \varphi \vdash \psi$ [Variant: $\varphi,\ \neg\psi \rightarrow \neg\varphi \vdash \psi$]

(.4) If $\Gamma,\ \neg\varphi,\ \neg\psi \vdash \varphi$, then $\Gamma, \neg\varphi \vdash \psi$ [Variant: $\neg\varphi,\ \neg\psi \rightarrow \varphi \vdash \psi$]

(.5) If $\Gamma,\ \varphi,\ \psi \vdash \neg\varphi$, then $\Gamma, \varphi \vdash \neg\psi$ [Variant: $\varphi,\ \psi \rightarrow \neg\varphi \vdash \neg\psi$]

(.6) If $\Gamma,\ \neg\varphi,\ \psi \vdash \varphi$, then $\Gamma, \neg\varphi \vdash \neg\psi$ [Variant: $\neg\varphi,\ \psi \rightarrow \varphi \vdash \neg\psi$]

(**88**) **Theorems:** Other Useful Tautologies. The foregoing allow us to more easily prove many classical and other useful theorems governing the classical connectives:

(.1) Some Basic Facts:

(.a) $(\varphi \rightarrow \psi) \equiv \neg(\varphi \,\&\, \neg\psi)$

(.b) $\neg(\varphi \rightarrow \psi) \equiv (\varphi \,\&\, \neg\psi)$

(.c) $(\varphi \rightarrow \psi) \equiv (\neg\varphi \vee \psi)$

(.2) Commutative and Associative Laws of &, $\vee$, and $\equiv$:

(.a) $(\varphi \,\&\, \psi) \equiv (\psi \,\&\, \varphi)$ (Commutativity of &)

(.b) $(\varphi \,\&\, (\psi \,\&\, \chi)) \equiv ((\varphi \,\&\, \psi) \,\&\, \chi)$ (Associativity of &)

(.c) $(\varphi \vee \psi) \equiv (\psi \vee \varphi)$ (Commutativity of $\vee$)

(.d) $(\varphi \vee (\psi \vee \chi)) \equiv ((\varphi \vee \psi) \vee \chi)$ (Associativity of $\vee$)

(.e) $(\varphi \equiv \psi) \equiv (\psi \equiv \varphi)$ (Commutativity of $\equiv$)

(.f) $(\varphi \equiv (\psi \equiv \chi)) \equiv ((\varphi \equiv \psi) \equiv \chi)$ (Associativity of $\equiv$)

(.3) Simple Biconditionals:

(.a) $\varphi \equiv \varphi$

(.b) $\varphi \equiv \neg\neg\varphi$

(.c) $\neg(\varphi \equiv \neg\varphi)$ (Form of Noncontradiction)

(.4) Conditionals and Biconditionals:

(.a) $(\varphi \rightarrow \psi) \rightarrow ((\psi \rightarrow \chi) \rightarrow (\varphi \rightarrow \chi))$

(.b) $(\varphi \equiv \psi) \equiv (\neg\varphi \equiv \neg\psi)$

(.c) $(\varphi \equiv \psi) \rightarrow ((\varphi \rightarrow \chi) \equiv (\psi \rightarrow \chi))$

(.d) $(\varphi \equiv \psi) \rightarrow ((\chi \rightarrow \varphi) \equiv (\chi \rightarrow \psi))$

(.e) $(\varphi \equiv \psi) \rightarrow ((\varphi \,\&\, \chi) \equiv (\psi \,\&\, \chi))$

(.f) $(\varphi \equiv \psi) \rightarrow ((\chi \,\&\, \varphi) \equiv (\chi \,\&\, \psi))$

(.g) $(\varphi \equiv \psi) \equiv ((\varphi \,\&\, \psi) \vee (\neg\varphi \,\&\, \neg\psi))$

(.h) $\neg(\varphi \equiv \psi) \equiv ((\varphi \,\&\, \neg\psi) \vee (\neg\varphi \,\&\, \psi))$

(.5) De Morgan's Laws:

(.a) $(\varphi \,\&\, \psi) \equiv \neg(\neg\varphi \vee \neg\psi)$

(.b) $(\varphi \vee \psi) \equiv \neg(\neg\varphi \,\&\, \neg\psi)$

(.c) $\neg(\varphi \,\&\, \psi) \equiv (\neg\varphi \vee \neg\psi)$

(.d) $\neg(\varphi \vee \psi) \equiv (\neg\varphi \,\&\, \neg\psi)$

(.6) Distribution Laws:

(.a) $(\varphi \,\&\, (\psi \vee \chi)) \equiv ((\varphi \,\&\, \psi) \vee (\varphi \,\&\, \chi))$

(.b) $(\varphi \vee (\psi \,\&\, \chi)) \equiv ((\varphi \vee \psi) \,\&\, (\varphi \vee \chi))$

(.7) Exportation and Importation:

(.a) $((\varphi \,\&\, \psi) \rightarrow \chi) \rightarrow (\varphi \rightarrow (\psi \rightarrow \chi))$ (Exportation)

(.b) $(\varphi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\varphi \,\&\, \psi) \rightarrow \chi)$ (Importation)

(.8) Other Miscellaneous Tautologies:

(.a) $(\varphi \to (\psi \to \chi)) \equiv (\psi \to (\varphi \to \chi))$ (Permutation)

(.b) $(\varphi \to \psi) \to ((\varphi \to \chi) \to (\varphi \to (\psi \,\&\, \chi)))$ (Composition)

(.c) $(\varphi \to \chi) \to ((\psi \to \chi) \to ((\varphi \vee \psi) \to \chi))$

(.d) $((\varphi \to \psi) \,\&\, (\chi \to \theta)) \to ((\varphi \,\&\, \chi) \to (\psi \,\&\, \theta))$ (Double Composition)

(.e) $((\varphi \,\&\, \psi) \equiv (\varphi \,\&\, \chi)) \equiv (\varphi \to (\psi \equiv \chi))$

(.f) $((\varphi \,\&\, \psi) \equiv (\chi \,\&\, \psi)) \equiv (\psi \to (\varphi \equiv \chi))$

(.g) $(\psi \equiv \chi) \to ((\varphi \vee \psi) \equiv (\varphi \vee \chi))$

(.h) $(\psi \equiv \chi) \to ((\psi \vee \varphi) \equiv (\chi \vee \varphi))$

(.i) $(\varphi \equiv (\psi \,\&\, \chi)) \to (\psi \to (\varphi \equiv \chi))$

We leave the proof of these tautologies, with the exception of (.8.i), as exercises.

(**89**) **Metarules/Derived Rules:** Reasoning with Biconditionals. Our standard axiomatization of negation and conditionalization and the standard definition of the $\equiv$ allow us to reason using all the *classical* introduction and elimination rules for the biconditional. However, we formulate them, in the first instance, as metarules.

(.1) Disjunctive Syllogism (alternative form):
If $\Gamma_1 \vdash \varphi \vee \psi$, $\Gamma_2 \vdash \varphi \equiv \chi$, and $\Gamma_3 \vdash \psi \equiv \theta$, then $\Gamma_1, \Gamma_2, \Gamma_3 \vdash \chi \vee \theta$
[Variant: $\varphi \vee \psi,\ \varphi \equiv \chi,\ \psi \equiv \theta \vdash \chi \vee \theta$]

(.2) $\equiv$Introduction ($\equiv$I):

If $\Gamma_1 \vdash \varphi \to \psi$ and $\Gamma_2 \vdash \psi \to \varphi$, then $\Gamma_1, \Gamma_2 \vdash \varphi \equiv \psi$
[Variant: $\varphi \to \psi,\ \psi \to \varphi \vdash \varphi \equiv \psi$]

(.3) $\equiv$Elimination ($\equiv$E) (Biconditional Syllogisms):

(.a) If $\Gamma_1 \vdash \varphi \equiv \psi$ and $\Gamma_2 \vdash \varphi$, then $\Gamma_1, \Gamma_2 \vdash \psi$ [Variant: $\varphi \equiv \psi,\ \varphi \vdash \psi$]

(.b) If $\Gamma_1 \vdash \varphi \equiv \psi$ and $\Gamma_2 \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \varphi$ [Variant: $\varphi \equiv \psi,\ \psi \vdash \varphi$]

(.c) If $\Gamma_1 \vdash \varphi \equiv \psi$ and $\Gamma_2 \vdash \neg\varphi$, then $\Gamma_1, \Gamma_2 \vdash \neg\psi$
[Variant: $\varphi \equiv \psi,\ \neg\varphi \vdash \neg\psi$]

(.d) If $\Gamma_1 \vdash \varphi \equiv \psi$ and $\Gamma_2 \vdash \neg\psi$, then $\Gamma_1, \Gamma_2 \vdash \neg\varphi$
[Variant: $\varphi \equiv \psi,\ \neg\psi \vdash \neg\varphi$]

(.e) If $\Gamma_1 \vdash \varphi \equiv \psi$ and $\Gamma_2 \vdash \psi \equiv \chi$, then $\Gamma_1, \Gamma_2 \vdash \varphi \equiv \chi$
[Variant: $\varphi \equiv \psi,\ \psi \equiv \chi \vdash \varphi \equiv \chi$]

(.f) If $\Gamma_1 \vdash \varphi \equiv \psi$ and $\Gamma_2 \vdash \varphi \equiv \chi$, then $\Gamma_1, \Gamma_2 \vdash \chi \equiv \psi$
[Variant: $\varphi \equiv \psi,\ \varphi \equiv \chi \vdash \chi \equiv \psi$]

In every case, the modally strict, $\vdash_\Box$-version of the rule holds and its justification requires only a trivial revision to the justification of the above versions. We leave the justification of these metarules, their variants, and their modally-strict versions and variants, as exercises and henceforth use the derived rules within proofs and derivations.

(**90**) **Metarule:** Derived Rule of Equivalence by Definition. Now that we have rules for reasoning with biconditionals, we can more easily derive, and make use of, the following rule:

(.1) **Rule of Equivalence by Definition** (Rule $\equiv$Df)
If $\varphi \equiv_{df} \psi$ is any definition-by-$\equiv$ and Γ is any premise set, then $\Gamma \vdash \varphi \equiv \psi$.

By convention (67.1), we omit the statement of the same rule for $\vdash_\Box$. But it should be kept in mind that facts of the form $\Gamma \vdash_\Box \varphi \equiv \psi$ derived from definitions by the $\vdash_\Box$ form of the rule play an essential role proving important theorems in what follows.

It should be noted here that once the Rule of Actualization (RA) in is formulated and justified in (135), one can strengthen Rule $\equiv$Df so that it asserts:

> If $\varphi \equiv_{df} \psi$ is a definition-by-$\equiv$, Γ is any premise set, and χ is any *closure* of $\varphi \equiv \psi$, then $\Gamma \vdash \chi$

Once Rule RA is in place then we can use the rules GEN, RN, and RA enough times to show that if $\varphi \equiv \psi$ is a theorem, then *any* closure of $\varphi \equiv \psi$ is a theorem. But this strengthened version of (.1) will not be explicitly formulated as such in what follows.

Now, given (.1), and the various rules for reasoning with biconditionals, (89.1) – (89.3), we can easily derive the rules that lets us directly infer a definiens from a definiendum and infer a definiendum from a definiens, from any instance of a definition-by-$\equiv$:

(.2) **Rule of Definiendum Elimination** (Rule $\equiv_{df}$**E**)
If $\Gamma \vdash \varphi$, then $\Gamma \vdash \psi$, whenever $\varphi \equiv_{df} \psi$ is a definition-by-$\equiv$.
[Variant: $\varphi \vdash \psi$, whenever ...]

(.3) **Rule of Definiendum Introduction** (Rule $\equiv_{df}$**I**)
If $\Gamma \vdash \psi$, then $\Gamma \vdash \varphi$, whenever $\varphi \equiv_{df} \psi$ is a definition-by-$\equiv$.
[Variant: $\psi \vdash \varphi$, whenever ...]

Again, by convention, we omit the statement of these rules for $\vdash_\Box$. Given that (.1) – (.3) hold for both $\vdash$ and $\vdash_\Box$, we may employ them in any reasoning environment without affecting the modal strictness, if any, of the reasoning. Once we gain some experience explicitly citing Rules $\equiv_{df}$E and $\equiv_{df}$I when proving theorems, we'll then revert to the classical citation 'by definition' when using these rules to draw conclusions from definitions-by-$\equiv$.

(**91**) **Metarule:** Conditions Permitting Biconditional Simplification. There is a situation that will occur frequently in which certain biconditionals can be simplified, namely, when one condition of the biconditional includes a number of conjuncts that are already known by hypothesis or by proof. The metarule for simplifying these biconditionals states that if there is a (modally-strict) derivation of $\varphi \equiv (\psi \,\&\, \chi)$ from Γ and there is a (modally-strict) derivation of ψ from Γ, then there is a (modally-strict) derivation of $\varphi \equiv \chi$ from Γ:

(.1) **Rule ≡S of Biconditional Simplification**
If $\Gamma \vdash \varphi \equiv (\psi \,\&\, \chi)$ and $\Gamma \vdash \psi$, then $\Gamma \vdash \varphi \equiv \chi$.
[Variant: $\varphi \equiv (\psi \,\&\, \chi), \psi \vdash \varphi \equiv \chi$]

The version for $\vdash_\Box$ is obtained by replacing $\vdash$ by $\vdash_\Box$. The Variant simply says we can infer (i.e., there is a derivation of) $\varphi \equiv \chi$ from the premises $\varphi \equiv (\psi \,\&\, \chi)$ and ψ. These rules are easily justified by (88.8.i).

A more general form of the above rule can also be derived. The Variant form states that from two assumptions, the first of which is a biconditional between φ and the conjunction of $\psi_1, \ldots, \psi_n$ ($n \geq 1$), and the second of which is just one of the ψ_i ($1 \leq i \leq n$), there is a (modally-strict) derivation of the equivalence that results by omitting the conjunct ψ_i from the right condition of the first assumption. Formally, for $1 \leq i \leq n$:

(.2) **Rule ≡S of Biconditional Simplification** (General Form)
If $\Gamma \vdash \varphi \equiv (\psi_1 \,\&\, \ldots \,\&\, \psi_n \,\&\, \chi)$ and $\Gamma \vdash \psi_i$,
then $\Gamma \vdash \varphi \equiv (\psi_1 \,\&\, \ldots \,\&\, \psi_{i-1} \,\&\, \psi_{i+1} \,\&\, \ldots \,\&\, \psi_n \,\&\, \chi)$

Variant:
$$\varphi \equiv (\psi_1 \,\&\, \ldots \,\&\, \psi_n \,\&\, \chi), \psi_i \vdash \varphi \equiv (\psi_1 \,\&\, \ldots \,\&\, \psi_{i-1} \,\&\, \psi_{i+1} \,\&\, \ldots \,\&\, \psi_n \,\&\, \chi)$$

We leave the justification as an exercise.

A special case of the Variant of (.2) applies to biconditionals derived from definitions-by-≡ in which the definiens has one or more existence claims as conjuncts. Consider any such instance of such a definition having the following form, in which the free occurrences of the variables $\alpha_1, \ldots, \alpha_n$ ($n \geq 1$) are interpreted under Convention (17.2):

$$\varphi(\alpha_1, \ldots, \alpha_n) \equiv_{df} \alpha_1\!\downarrow \,\&\, \ldots \,\&\, \alpha_n\!\downarrow \,\&\, \chi$$

Now let $\tau_1, \ldots, \tau_n$ be any terms substitutable for $\alpha_1, \ldots, \alpha_n$. Then where $1 \leq i \leq n$, we have the following special case of the general form of Rule ≡S:

$$\varphi^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n} \equiv (\tau_1\!\downarrow \,\&\, \ldots \,\&\, \,\&\, \tau_n\!\downarrow \,\&\, \chi^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}), \tau_i\!\downarrow \vdash$$
$$\varphi^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n} \equiv (\tau_1\!\downarrow \,\&\, \ldots \,\&\, \tau_{i-1}\!\downarrow \,\&\, \tau_{i+1}\!\downarrow \,\&\, \ldots \,\&\, \tau_n\!\downarrow \,\&\, \chi^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n})$$

This special case will often be applied in reasoning from definitions-by-≡ in which the definiens contains a non-zero number of existence clauses as conjuncts.

(**92**) **Remark:** Not All Tautologies Are Yet Derivable. Rule MP and our axioms (38.1) – (38.3) for negations and conditionals are not yet sufficient for deriving all of the formulas that qualify as tautologies, as the latter notion was defined in Section 6.2. We discovered in that section that our system contains a new class of tautologies that arise in connection with 0-ary relation terms of the form $[\lambda\,\varphi]$. Instances of the following schemata are members of this new class of tautologies: $[\lambda\,\varphi] \rightarrow \varphi$, $[\lambda\,\varphi] \equiv \varphi$, $[\lambda\,\varphi] \rightarrow \neg\neg\varphi$, etc. To derive these tautologies, we must first prove that $[\lambda\,\varphi] \equiv \varphi$ is a theorem (111.2), and to do that, we will need to show that $[\lambda\,\varphi] = \varphi$ is a theorem (111.1). The derivations of these latter theorems appeal to η-Conversion, GEN, Rule ∀E (a rule of quantification theory derived in item (93) below), and Rule =E (i.e., the rule for the substitution of identicals, derived in item (110) below). Once we've derived all of these key principles, and (111.2) in particular, we will be in a position to prove Metatheorem ⟨9.2⟩, i.e., that all tautologies are derivable. This Metatheorem is proved in the Appendix to this chapter. With such a result, we can derive Rule T, which is formulable using the semantic notions defined in Section 6.2, as a rule for our system:

> **Rule T**
> If $\Gamma \vdash \varphi_1$ and … and $\Gamma \vdash \varphi_n$, then if $\{\varphi_1, \ldots, \varphi_n\}$ tautologically implies ψ, then $\Gamma \vdash \psi$.

Rule T asserts that ψ is derivable from Γ whenever the formulas of which it is a tautological consequence are all derivable from Γ. We won't use this rule in proving theorems, since it requires semantic notions. But it is a valid shortcut. Rule T is proved as Metatheorem ⟨9.4⟩ in the Appendix to this chapter.

9.6 The Theory of Quantification

(**93**) **Metarules/Derived Rules:** ∀Elimination (∀E). The elimination rule for the universal quantifier has two forms (with the first being the primary form):

> **Rule ∀E**
> (.1) If $\Gamma_1 \vdash \forall\alpha\varphi$ and $\Gamma_2 \vdash \tau{\downarrow}$, then $\Gamma_1, \Gamma_2 \vdash \varphi^{\tau}_{\alpha}$, provided τ is substitutable for α in φ [Variant: $\forall\alpha\varphi,\ \tau{\downarrow} \vdash \varphi^{\tau}_{\alpha}$, if τ is substitutable for α]
>
> (.2) If $\Gamma \vdash \forall\alpha\varphi$, then $\Gamma \vdash \varphi^{\tau}_{\alpha}$, provided τ is (a) a primitive constant, a variable, or a core λ-expression, and (b) substitutable for α in φ
> [Variant: $\forall\alpha\varphi \vdash \varphi^{\tau}_{\alpha}$, provided …]

(Here, and in what follows, the ellipsis should be filled in by the provisos of the official form of the rule.) In the usual manner, we may conceive the variants as derived rules and use them to produce genuine derivations.

Rule ∀E and its Variant have special cases when τ is the variable α. Since every variable α is substitutable for itself in any formula φ (see Metatheorem ⟨7.4⟩ in (15)) and $\varphi^{\alpha}_{\alpha} = \varphi$ (see Metatheorem ⟨7.2⟩ in (14)), the following special case obtains:

Rule ∀E Special Case
(.3) If $\Gamma \vdash \forall\alpha\varphi$, then $\Gamma \vdash \varphi$ [Variant: $\forall\alpha\varphi \vdash \varphi$]

(**94**) **Remark:** A Misuse of Rule ∀E. Note that the following attempt to derive a contradiction involves a misuse of Rule ∀E:

Let φ be the formula $\neg yF$ and formulate the following instance of the Comprehension Principle for Abstract Objects: $\exists x(A!x \,\&\, \forall F(xF \equiv \neg yF))$. So by GEN, we may derive as a theorem:

(ϑ) $\forall y \exists x(A!x \,\&\, \forall F(xF \equiv \neg yF))$

Hence, by Rule ∀E (93.2) [Variant], it follows that:

(ξ) $\exists x(A!x \,\&\, \forall F(xF \equiv \neg xF))$

From this, we can derive a contradiction.[142]

The problem with this bit of reasoning is the invalid application of Rule ∀E (93.2) [Variant] to infer (ξ) from (ϑ). (ϑ) has the form $\forall y\psi$ and (ξ) has the form ψ^x_y, and so the inference from (ϑ) to (ξ) doesn't obey the condition that x must be substitutable for y in the formula ψ.[143]

(**95**) **Theorems:** Classical Quantifier Axioms as Theorems. Our system now yields two quantification principles as theorems:

(.1) $\forall\alpha\varphi \to \varphi^{\tau}_{\alpha}$, provided τ is (a) a primitive constant, a variable, or a core λ-expression, and (b) substitutable for α in φ.

(.2) $\forall\alpha(\varphi \to \psi) \to (\varphi \to \forall\alpha\psi)$, provided α is not free in φ

Versions of these theorems are often used as the principal axioms of classical quantification theory, with GEN as a primitive rule of inference. Cf. Mendelson

[142] We can see this more easily if we take the liberty of reasoning with a rule of inference we haven't officially introduced yet, namely Rule ∃E (102). For assume a is an arbitrary such object, so that we know $A!a \,\&\, \forall F(aF \equiv \neg aF)$. By &E, it follows that $\forall F(aF \equiv \neg aF)$. Hence by a special case of Rule ∀E (93.3), it follows that $aF \equiv \neg aF$, which by (88.3.c) is a contradiction.

[143] In this case, ψ is the formula $\exists x(A!x \,\&\, \forall F(xF \equiv \neg yF))$. Then the definition of *substitutable for* in (15) requires that for x to be substitutable for y in ψ, every variable that occurs free in the term x must remain free after we substitute x for y in ψ. But x itself, which is free in the term x, doesn't remain free when x is substituted for y in ψ to produce ψ^x_y, i.e., $\exists x(A!x \,\&\, \forall F(xF \equiv \neg xF))$. Instead, x is captured by (i.e., falls within the scope of) the existential quantifier $\exists x$ in ψ^x_y.

1964 [1997, 69, axioms (A4) and (A5)]; (A4) is a simpler version of (.1) because Mendelson's formulation of predicate logic (with function terms) assumes that all terms have a denotation.

Clearly, since $\varphi^{\alpha}_{\alpha}$ is just φ, formulas of the form:

(.3) $\forall\alpha\varphi \to \varphi$

are special cases of (.1).

(**96**) **Metarule/Derived Rule:** Rule ∀I (Universal Introduction) or Generalization on Constants. We introduce and explain the Rule ∀I by way of an example. We often argue as follows:

> Let P be an arbitrary property and a an arbitrary object; then as an instance of the tautology $\varphi \equiv \varphi$ (88.3.a), we know $Pa \equiv Pa$. Since a is arbitrary, the biconditional holds for all objects, i.e., $\forall x(Px \equiv Px)$. Since P is arbitrary, this last claim holds for all properties, i.e., $\forall F\forall x(Fx \equiv Fx)$.

The last two steps of this reasoning are justified by Rule ∀I, which allows us to universally generalize once we have reached a conclusion about arbitrarily chosen entities, as long as we haven't invoked any special assumptions about the chosen entities. In the example we just gave, no special facts about the property P or the individual a played a part in our intermediate conclusion that $Pa \equiv Pa$.

Of course, the theorem $\forall F\forall x(Fx \equiv Fx)$ could have been established without appealing to Rule ∀I, as follows: $Fx \equiv Fx$ is an instance of the tautology $\varphi \equiv \varphi$ (88.3.a) and so by two applications of GEN, it follows that $\forall F\forall x(Fx \equiv Fx)$. But though we can reason using variables in this way, it is nevertheless sometimes helpful (indeed, clearer) to use arbitrarily chosen (primitive) constants instead of free variables when reasoning.

To formulate Rule ∀I generally, we first introduce some notation. Where τ is any constant and α any variable of the same type as τ:

- φ^{α}_{τ} is the result of replacing every occurrence of the constant τ in φ by an occurrence of α

We then have:

> **Rule ∀I**
> If $\Gamma \vdash \varphi$ and τ is a primitive constant that does not occur in Γ or $\mathbf{\Lambda}$, then $\Gamma \vdash \forall\alpha\varphi^{\alpha}_{\tau}$, provided α is a variable that does not occur in φ.
> [Variant: $\varphi \vdash \forall\alpha\varphi^{\alpha}_{\tau}$, provided ...]

Note that only one distinguished constant, namely $E!$, is used in the statement of the axioms in $\mathbf{\Lambda}$ and so $E!$ is not an acceptable value for τ in Rule ∀I. So, for example, from axiom (39.4), i.e., $\Diamond\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$, we cannot conclude $\forall F(\Diamond\exists x(Fx \,\&\, \neg\mathcal{A}Fx)$ by Rule $\forall$I.

Here is another example of how we will use Rule $\forall$I. Consider the following reasoning that shows $\forall x(Px \to Qx), \forall yPy \vdash \forall xQx$:

1.	$\forall x(Px \to Qx)$	Premise
2.	$\forall yPy$	Premise
3.	$Pa \to Qa$	from 1, by $\forall$E (93.2) [Variant]
4.	Pa	from 2, by $\forall$E (93.2) [Variant]
5.	Qa	from 3, 4, by MP
6.	$\forall x(Px \to Qx), \forall yPy \vdash \forall xQx$	from 1–5, by $\forall$I

In this example, we set $\Gamma = \{\forall x(Px \to Qx), \forall yPy\}$, $\varphi = Qa$, and $\tau = a$. Given that $\forall$E is a derived rule as well as a metarule, lines 1–5 constitute a genuine derivation that is a witness to $\Gamma \vdash \varphi$. Since a doesn't occur in Γ or $\mathbf{\Lambda}$, and x doesn't occur in φ, we have an instance of the Rule $\forall$I in which α is the variable x, which we can then apply to lines 1–5 to infer the derivability claim on line 6.

Since $\forall$I is a metarule, we could have reached the conclusion on line 6 without it using the following reasoning, which doesn't involve the constant a:

1.	$\forall x(Px \to Qx)$	Premise
2.	$\forall yPy$	Premise
3.	$Px \to Qx$	from 1, by $\forall$E (93.3)
4.	Px	from 2, by $\forall$E (93.2) [Variant]
5.	Qx	from 3,4, by MP
6.	$\forall x(Px \to Qx),\ \forall yPy\ \vdash\ \forall xQx$	from 1–5, by GEN

The application of GEN on line 6 is legitimate since we have correctly derived Qx on line 5 from the premises $\forall x(Px \to Qx)$ and $\forall yPy$ and the variable x doesn't occur free in the premises. Of course, GEN itself is a metarule, but we already know how to eliminate it.

(**97**) **Lemmas:** Re-replacement Lemmas. In the following re-replacement lemmas, we assume that α and β are variables of the same type as term τ:

(.1) If β is substitutable for α in φ and β doesn't occur free in φ, then α is substitutable for β in φ^{β}_{α} and $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta} = \varphi$.

(.2) If τ is a constant symbol that doesn't occur in φ, then $(\varphi^{\tau}_{\alpha})^{\beta}_{\tau} = \varphi^{\beta}_{\alpha}$.

(.3) If β is substitutable for α in φ and doesn't occur free in φ, and τ is any term substitutable for α in φ, then $(\varphi^{\beta}_{\alpha})^{\tau}_{\beta} = \varphi^{\tau}_{\alpha}$.

It may help to read the following Remark before attempting to prove the above.

(**98**) **Remark:** Explanation of the Re-replacement Lemmas. By discussing (97.1) in some detail, (97.2) and (97.3) become more transparent and less in need of commentary. It is relatively easy to show that, in general, in the absence of any preconditions, $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta} \neq \varphi$. The variable α may occur in $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$ at a place where it does not occur in φ, or α may occur in φ at a place where it does not occur in $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$. Here is an example of each case:

- $\varphi = Ryx$. Then $\varphi^{y}_{x} = Ryy$ and though x is substitutable for y in φ^{y}_{x} (i.e., y doesn't fall under the scope of any variable-binding operator that binds x), $(\varphi^{y}_{x})^{x}_{y} = Rxx$. Hence $(\varphi^{y}_{x})^{x}_{y} \neq \varphi$. In this example, x occurs at a place in $(\varphi^{y}_{x})^{x}_{y}$ where it does not occur in φ.

- $\varphi = \forall yRxy$. Then $\varphi^{y}_{x} = \forall yRyy$. Since x is trivially substitutable for y in φ^{y}_{x} (there are no free occurrences of y in φ^{y}_{x}), $(\varphi^{y}_{x})^{x}_{y} = \varphi^{y}_{x} = \forall yRyy$. By inspection, then, $(\varphi^{y}_{x})^{x}_{y} \neq \varphi$. In this example, x occurs at a place in φ where it does not occur in $(\varphi^{y}_{x})^{x}_{y}$.

These two examples nicely demonstrate why the two antecedents of (97.1) are crucial. The first example fails the proviso that y not occur free in Ryx; the second example fails the proviso that y be substitutable for x in $\forall yRxy$. But here is an example of (97.1) in which the antecedents obtain:

- $\varphi = \forall yPy \to Qx$. In this example, the free occurrence of x is not within the scope of the quantifier $\forall y$. So y is substitutable for x in φ and y does not occur free in φ. Thus, $\varphi^{y}_{x} = \forall yPy \to Qy$, and since y has a free occurrence in φ^{y}_{x} not under the scope of a variable-binding operator binding x, x is substitutable for y in φ^{y}_{x}. Hence $(\varphi^{y}_{x})^{x}_{y} = \forall yPy \to Qx$, and so $(\varphi^{y}_{x})^{x}_{y} = \varphi$.

These remarks and the proof of (97.1) should suffice to clarify the remaining two replacement lemmas. (97.1) is used to prove (99.13), (103.10), and the Rule of Alphabetic Variants (114). Lemma (97.3) is used in the proof of (103.7).

(**99**) **Theorems:** Basic Theorems of Quantification Theory. The following are all basic consequences of our quantifier axioms and (derived) rules:

(.1) $\forall\alpha\forall\beta\varphi \equiv \forall\beta\forall\alpha\varphi$

(.2) $\forall\alpha(\varphi \equiv \psi) \equiv (\forall\alpha(\varphi \to \psi) \;\&\; \forall\alpha(\psi \to \varphi))$

(.3) $\forall\alpha(\varphi \equiv \psi) \to (\forall\alpha\varphi \equiv \forall\alpha\psi)$

(.4) $\forall\alpha(\varphi \;\&\; \psi) \equiv (\forall\alpha\varphi \;\&\; \forall\alpha\psi)$

(.5) $\forall\alpha_1 \ldots \forall\alpha_n\varphi \to \varphi$

(.6) $\forall\alpha\forall\alpha\varphi \equiv \forall\alpha\varphi$

(.7) $(\varphi \to \forall\alpha\psi) \equiv \forall\alpha(\varphi \to \psi)$, provided α is not free in φ

(.8) $(\forall\alpha\varphi \vee \forall\alpha\psi) \to \forall\alpha(\varphi \vee \psi)$

(.9) $(\forall\alpha(\varphi \to \psi) \,\&\, \forall\alpha(\psi \to \chi)) \to \forall\alpha(\varphi \to \chi)$

(.10) $(\forall\alpha(\varphi \equiv \psi) \,\&\, \forall\alpha(\psi \equiv \chi)) \to \forall\alpha(\varphi \equiv \chi)$

(.11) $\forall\alpha(\varphi \equiv \psi) \equiv \forall\alpha(\psi \equiv \varphi)$

(.12) $\forall\alpha\varphi \to \forall\alpha(\psi \to \varphi)$

(.13) $\forall\alpha\varphi \equiv \forall\beta(\varphi^{\beta}_{\alpha})$,
provided β is substitutable for α in φ and doesn't occur free in φ

(.14) $\forall\alpha_1 \ldots \forall\alpha_n(\varphi \to \psi) \to (\forall\alpha_1 \ldots \forall\alpha_n\varphi \to \forall\alpha_1 \ldots \forall\alpha_n\psi)$,
for any distinct variables $\alpha_1, \ldots, \alpha_n$ and $n \geq 2$

(.15) $\forall\alpha_1 \ldots \forall\alpha_n(\varphi \to \psi) \to (\varphi \to \forall\alpha_1 \ldots \forall\alpha_n\psi)$,
provided $\alpha_1, \ldots, \alpha_n$ ($n \geq 2$) are distinct variables that don't occur free in φ

The two provisos on (.13) can be explained by referencing and adapting the examples used in Remark (98) that helped us to understand the antecedent of the Re-replacement Lemma (97.1):

- In the formula $\varphi = Rxy$, y is substitutable for x but also occurs free. Without the second proviso in (.13), we could set α to x and β to y and obtain the instance: $\forall xRxy \equiv \forall yRyy$. Clearly, this is not valid: the left side asserts that everything bears R to y while the right asserts that everything bears R to itself.

- In the formula $\varphi = \forall yRxy$, y is not substitutable for x and does not occur free. Without the first proviso in (.13), we could set α to x and β to y and obtain the instance: $\forall x\forall yRxy \equiv \forall y\forall yRyy$. Again, clearly, this is not valid: the left side is true when everything bears R to everything while the right side, which by (.6) is equivalent to $\forall yRyy$, is true only when everything bears R to itself.

(.13) is a special case of the interderivability of alphabetic variants; indeed, it is a special case of a special case. The interderivability of alphabetically-variant universal generalizations is a special case of the interderivability of alphabetically-variant formulas of arbitrary complexity. Note that there are two basic ways in which a universal generalization of the form $\forall\alpha\varphi$ can have an alphabetic variant. (.13) concerns one of those ways, namely, alphabetic variants of the form $\forall\beta(\varphi^{\beta}_{\alpha})$. But $\forall\alpha\varphi$ can also have alphabetic variants of the form $\forall\alpha(\varphi')$, where φ' is an alphabetic variant of φ. We aren't yet in a position to prove the interderivability of $\forall\alpha\varphi$ and $\forall\alpha(\varphi')$ in such a case, much less prove the interderivability of alphabetically-variant formulas of *arbitrary* complexity. (.13) tells us only that whenever we have established a theorem of the form $\forall\alpha\varphi$, we may infer any formula with the same exact form but which differs throughout only by the choice of the variable bound by the leftmost universal quantifier, provided the choice of the new variable is a safe one, i.e., one that will preserve the meaning of the original formula when the substitution is carried out.

(**100**) **Metarule/Derived Rule:** Corollary to Rule $\forall$I. Using fact (99.13), we may prove the following:

> **Corollary to Rule $\forall$I**:
> If $\Gamma \vdash \varphi^{\tau}_{\alpha}$, and τ is a primitive constant (of the same type as α) that doesn't occur in Γ, $\mathbf{\Lambda}$, or φ, then $\Gamma \vdash \forall\alpha\varphi$ and there is a derivation of $\forall\alpha\varphi$ from Γ in which τ doesn't occur. [Variant: $\varphi^{\tau}_{\alpha} \vdash \forall\alpha\varphi$, provided ...]

In the usual manner, the ellipsis in the Variant states the conditions included in the official version of the rule.

(**101**) **Metarules/Derived Rules:** $\exists$Introduction ($\exists$I). The metarules of $\exists$Introduction allow us to infer the existence of derivations of existential generalizations, though their variant forms yield derived rules that let us existentially generalize, within a derivation, on any term τ known to have a denotation. Rule $\exists$I has two forms: one that applies to any term whatsoever and a restricted form that applies to terms whose significance is axiomatic:

> **Rule $\exists$I**
>
> (.1) If $\Gamma_1 \vdash \varphi^{\tau}_{\alpha}$ and $\Gamma_2 \vdash \tau\!\downarrow$, then $\Gamma_1,\Gamma_2 \vdash \exists\alpha\varphi$, provided τ is substitutable for α in φ
> [Variant: $\varphi^{\tau}_{\alpha}, \tau\!\downarrow \vdash \exists\alpha\varphi$, provided τ is substitutable for α in φ]
>
> (.2) If $\Gamma \vdash \varphi^{\tau}_{\alpha}$, then $\Gamma \vdash \exists\alpha\varphi$, provided τ is (a) a primitive constant, a variable, or a core λ-expression, and (b) substitutable for α in φ. [Variant: $\varphi^{\tau}_{\alpha} \vdash \exists\alpha\varphi$, provided ...]

Two examples of the Variant version of (.2) are: $Gy \vdash \exists xGx$ and $Gy \vdash \exists F Fy$. In the first case, φ is Gx, α is x, τ is y, φ^{τ}_{α} is Gy. So the conditions are met: y is a variable and substitutable for x in Gx. In the second case, φ is Fy, α is F, τ is G, φ^{τ}_{α} is Gy. So the conditions are met: G is a variable and substitutable for F in Fy. Note that even $Gx \vdash \exists xGx$ and $Fy \vdash \exists F Fy$ are also instances of the Variant version.

However, $Pa \,\&\, Qx \vdash \exists x(Px \,\&\, Qx)$ is *not* a valid instance of the rule. Here, φ is $Px \,\&\, Qx$ and the premise $Pa \,\&\, Qx$ does not have the required form φ^{a}_{x}, since that is defined to be the formula $Pa \,\&\, Qa$, by (14). Thus, we can't existentially generalize on a in $Pa \,\&\, Qx$ by using the quantifier $\exists x$ to conclude $\exists x(Px \,\&\, Qx)$, for that would invalidly capture the free variable x. Similarly, from $Pa \,\&\, Fb$ we may not validly use $\exists$I to infer $\exists F(Fa \,\&\, Fb)$.

Note also that the inference from $\neg P\iota xQx$ and $\iota xQx{\downarrow}$ to $\exists x\neg Px$ is justified by the Variant version of (.1), as is the inference from $\neg[\lambda x\, \psi]y$ and $[\lambda x\, \psi]{\downarrow}$ to $\exists F\neg Fy$. Since versions of Rule $\exists$I are covered in detail in basic courses on predicate logic, we omit further examples and explanation of the conditions that must be satisfied for the rule to be applied.

(**102**) **Metarule:** $\exists$Elimination ($\exists$E) on Constants. Let τ be a primitive constant (i.e., not one introduced by a definition). If we validly reason our way to ψ from φ^{τ}_{α} (and some other premises) without making any special assumptions about τ other than φ^{τ}_{α}, then $\exists$E allows us to discharge our assumption about τ and validly conclude that we can derive ψ from (any premises we used and) $\exists\alpha\varphi$:

> **Rule $\exists$E**
> If $\Gamma, \varphi^{\tau}_{\alpha} \vdash \psi$, then $\Gamma, \exists\alpha\varphi \vdash \psi$, provided τ is a primitive constant that does not occur in φ, ψ, Γ, or $\mathbf{\Lambda}$.

We will often use this rule along with conditional proof when proving a theorem of the form $\exists\alpha\varphi \rightarrow \psi$, as follows. First, we assume $\exists\alpha\varphi$, for conditional proof. Then we assume that τ is an arbitrary such φ, i.e., that φ^{τ}_{α}, where τ is a *fresh*, primitive constant (of the same type as α) that hasn't previously appeared in the context of reasoning or in our axioms. Then, once we derive ψ from φ^{τ}_{α}, we appeal to Rule $\exists$E to discharge the assumption φ^{τ}_{α} and conclude that we've derived ψ only from the assumption that $\exists\alpha\varphi$. Then we conclude $\exists\alpha\varphi \rightarrow \psi$ by conditional proof.

(**103**) **Theorems:** Further Theorems of Quantification Theory. The various rules for quantification theory introduced thus far facilitate the derivation of many of the following theorems:

(.1) $\forall\alpha\varphi \rightarrow \exists\alpha\varphi$

(.2) $\neg\forall\alpha\varphi \equiv \exists\alpha\neg\varphi$

(.3) $\forall\alpha\varphi \equiv \neg\exists\alpha\neg\varphi$

(.4) $\neg\exists\alpha\varphi \equiv \forall\alpha\neg\varphi$

(.5) $\exists\alpha(\varphi \,\&\, \psi) \rightarrow (\exists\alpha\varphi \,\&\, \exists\alpha\psi)$

(.6) $\exists\alpha(\varphi \vee \psi) \equiv (\exists\alpha\varphi \vee \exists\alpha\psi)$

(.7) $\exists\alpha\varphi \equiv \exists\beta(\varphi^{\beta}_{\alpha})$,
provided β is substitutable for α in φ and doesn't occur free in φ

(.8) $(\forall\alpha\varphi \,\&\, \forall\alpha\psi) \rightarrow \forall\alpha(\varphi \equiv \psi)$

(.9) $(\neg\exists\alpha\varphi \,\&\, \neg\exists\alpha\psi) \rightarrow \forall\alpha(\varphi \equiv \psi)$

(.10) $(\exists\alpha\varphi \,\&\, \neg\exists\alpha\psi) \rightarrow \neg\forall\alpha(\varphi \equiv \psi)$

(.11) $\exists\alpha\exists\beta\varphi \equiv \exists\beta\exists\alpha\varphi$

These are all classical theorems of predicate logic.

9.7 Logical Existence, Identity, and Truth

(**104**) **Theorems:** Every 0-ary Relation Term and Formula is Significant. For any relation term Π^0, it is a theorem that (.1) Π^0 exists; and for any formula φ, it is a theorem that (.2) φ exists:

(.1) $\Pi^0\!\downarrow$, for any 0-ary relation term Π^0

(.2) $\varphi\!\downarrow$, for any formula φ

Intuitively, (.1) and (.2) tells us, respectively, that every 0-ary relation term and every formula of our language has a denotation. But though every formula φ provably has a denotation, it doesn't follow that every λ-expression of the form $[\lambda\nu\,\varphi]$ constructed from φ has a denotation. In particular, if φ has a free occurrence of ν, then although φ is significant, $[\lambda\nu\varphi]$ may fail to be significant. For example, the formula $\exists F(xF \,\&\, \neg Fx)$, in which x occurs free, is significant, by (.2), but $[\lambda x\, \exists F(xF \,\&\, \neg Fx)]$ provably fails to be significant; this is the λ-expression that leads to the Clark/Boolos paradox.[144]

[144] Thus, while there is, for every formula φ, a corresponding term of the form $[\lambda\nu\,\varphi]$, it doesn't follow that the corresponding term has a denotation. So the argument in Oppenheimer & Zalta 2011 is still valid, though it needs to be reframed. The argument in that paper, to the conclusion that object theory can't be straightforwardly represented in functional type theory (FTT), rested on two premises:

(a) FTT semantically analyzes a quantifier claim such as $\forall\nu\varphi$, where ν occurs free in φ and so bound by $\forall\nu$, by first converting φ to a term of the form $[\lambda\nu\,\varphi]$, and

(b) object theory includes some formulas φ and variables ν that can't be converted to the corresponding term $[\lambda\nu\,\varphi]$, on pain of contradiction. (For example, where φ is $\exists F(xF \,\&\, \neg Fx)$ and ν is x, $[\lambda x\, \exists F(xF \,\&\, \neg Fx)]$ was not even well-formed.)

In the formalism we used in 2011, a λ-expression that included *any* encoding subformulas in its matrix failed to be well-formed.

But in the present work, the expression $[\lambda\nu\,\varphi]$ is a *well-formed* term, for every formula φ, even if occurrences of ν bound by the λ occur in encoding position in φ. Of course, not every φ gives rise to a *significant* term of the form $[\lambda\nu\,\varphi]$. So the argument in Oppenheimer & Zalta 2011 has to

(**105**) **Remark:** An Interesting Fact Made Clear by the Foregoing. It is worth remarking on how theorem (104.2) implies an important difference between what can be inferred from the *truth* of a formula and what can be inferred from the *existence* of the proposition signified by such a formula. Consider, for example, an encoding formula of the form $\imath y\varphi[\lambda z\,\psi]$. By axiom (39.5.b), the *truth* of $\imath y\varphi[\lambda z\,\psi]$ implies that both $\imath y\varphi\downarrow$ and $[\lambda z\,\psi]\downarrow$. So by the contrapositive, if either $\neg\imath y\varphi\downarrow$ or $\neg[\lambda z\,\psi]\downarrow$, then $\neg\imath y\varphi[\lambda z\,\psi]$.

However, the claim $(\imath y\varphi[\lambda z\,\psi])\downarrow$, i.e., that the proposition $\imath y\varphi[\lambda z\,\psi]$ *exists*, does *not* imply that either $\imath y\varphi\downarrow$ or $[\lambda z\,\psi]\downarrow$. One can find φ and ψ such that all three of the following hold: $\imath y\varphi[\lambda z\,\psi]\downarrow$, $\neg\imath y\varphi\downarrow$, and $\neg[\lambda z\,\psi]\downarrow$. For let $\imath y\varphi$ be some *closed* description that couldn't be significant, say $\imath y(Py\&\neg Py)$, and let $[\lambda z\psi]$ be some *closed* λ-expression such as the one involved in the Clark/Boolos paradox, so that we know $\neg[\lambda z\psi]\downarrow$. It is still a theorem that $(\imath y\varphi[\lambda z\psi])\downarrow$, as the proof of (104.2) shows: for this particular φ and ψ, $[\lambda x\,\imath y\varphi[\lambda z\psi]]\downarrow$ is axiomatic, by (39.2) – since the initial λ doesn't bind x anywhere in the matrix, it is a core λ-expression (9.2); no variable bound by the λ occurs in formula a primary term in encoding position in the matrix (9.1). So it follows by definition (20.3) that $(\imath y\varphi[\lambda z\,\psi])\downarrow$.

Of course, the proposition asserted to exist is necessarily false. By hypothesis, $\neg\imath y\varphi\downarrow$ and $\neg[\lambda z\,\psi]\downarrow$, and each of these facts independently yield $\neg\imath y\varphi[\lambda z\psi]$, by (39.5.b). Since this derivation is modally strict, it follows by RN that $\Box\neg\imath y\varphi[\lambda z\,\psi]$. This makes it clear that although the proposition $\imath y\varphi[\lambda z\,\psi]$ exists, it is necessarily false, given that its primary terms fail to denote.

In general, every formula has truth conditions and denotes a proposition, but that fact doesn't imply that the terms in the formula denote something.[145] In particular, $\imath y\varphi[\lambda z\,\psi]\downarrow$ is a theorem (indeed, an instance of (104.2)), whether or not the terms $\imath y\varphi$ and $[\lambda z\,\psi]$ are significant.

(**106**) **Theorems:** The Necessity of Logical Existence. It is derivable, for any term τ, that if τ exists, then necessarily τ exists:

$$\tau\downarrow \to \Box\tau\downarrow$$

be expressed a bit differently. In particular, premise (b) in the argument of Oppenheimer & Zalta 2011 has to be revised to assert that object theory includes some formulas φ in which ν occurs free in encoding position but which can't be converted to a *denoting* term of the form $[\lambda\nu\,\varphi]$, on pain of contradiction. This is sufficient to make the point in Oppenheimer & Zalta 2011, since FTT's analysis of quantification requires that every formula φ in which the variable ν is free be convertible to a term of the form $[\lambda\nu\,\varphi]$ that has a denotation.

[145] We'll see an exception to this when we move to type theory in Chapter 15, where the increased expressive power allows us to form definite descriptions that are also formulas. In particular, we'll be able to form the following description of a proposition: $\imath p(p\,\&\,\neg p)$ ("the proposition that is both true and not true"). This expression is both a term and a formula, and when in formula position, it asserts: the proposition that is both true and not true is true. So this is a formula that has truth conditions (clearly, it is false) but which does *not* denote a proposition.

Thus, when any individual or relation exists in the logical sense, it necessarily does so and we can reference it in any modal context. Note that when τ is an individual term, this theorem doesn't imply that if τ is concrete, then τ is necessarily concrete. So-called contingent beings are ones that are concrete in some worlds but not others. They aren't contingent because they go in and out of logical existence from world to world. It is also important to note that we shall derive the necessity of logical non-existence in (169.3), once we have derived the principles of modal logic needed to prove it.

(**107**) **Theorems:** Identity Implies Existence. For any terms τ and σ, if $\tau = \sigma$, then both τ and σ exist:

(.1) $\tau = \sigma \rightarrow \tau\downarrow$

(.2) $\tau = \sigma \rightarrow \sigma\downarrow$

Note that this theorem holds even when τ or σ is a definite description or a λ-expression. If a term τ appears in a true identity statement, then it has a denotation. Similarly, if we can establish either $\neg\tau\downarrow$ or $\neg\sigma\downarrow$, then the contrapositive of (.1) or (.2) yields $\neg(\tau = \sigma)$. In such a case, we can infer $\tau \neq \sigma$, by definition (24).

(**108**) **Theorems:** Identical Relations are Necessarily Equivalent. It is now easily provable, for any n-ary relation terms ($n \geq 0$), that (.1) if an identity holds when those terms flank the identity sign, then the relations they signify are necessarily equivalent:

(.1) $\Pi = \Pi' \rightarrow \Box\forall x_1 \ldots \forall x_n(\Pi x_1 \ldots x_n \equiv \Pi' x_1 \ldots x_n)$,
where Π and Π' are any n-ary relation terms ($n \geq 0$) in which $x_1, \ldots, x_n$ don't occur free

The converse, however, does not hold in the present system.

Since formulas are 0-ary relation terms, it is a 0-ary instance of the above that (.2) if φ is identical to ψ, then necessarily, φ if and only ψ, where φ and ψ are any formulas:

(.2) $\varphi = \psi \rightarrow \Box(\varphi \equiv \psi)$

We formulate these theorems not only because they capture intuitions that should be preserved, but also to make the following observation.

(**109**) **Remark:** Definitions by Equivalence vs. Definition by Identity. Theorems (104.2) and (108.2) have an interesting consequence for our theory of definitions, namely, that one *could* eliminate definitions-by-equivalence in favor of definitions-by-identity. To see why, note that formulas are terms and so a definition of the form $\varphi =_{df} \psi$ is a definition-by-identity of the form $\tau =_{df} \sigma$.

So by (73), the inferential role of $\varphi =_{df} \psi$ is to introduce (the closures of) $(\psi\downarrow \to (\varphi = \psi))\ \&\ (\neg\psi\downarrow \to \neg\varphi\downarrow)$ as necessary axioms. But if every formula ψ is significant (104.2), then one can derive $\varphi = \psi$ from any definition-by-identity $\varphi =_{df} \psi$. So, by (108.2), one can derive $\Box(\varphi \equiv \psi)$ from such definitions. Since one can derive necessarily equivalences from definitions-by-identity, our system offers the option of foregoing definitions-by-equivalence in favor of definitions-by-identity.[146]

However, we shall not eschew definitions-by-equivalence in favor of definitions-by-identity. The reason has to do with maximizing our system's ability to preserve the hyperintensional contexts of natural language when the sentences of natural language are represented within our system. For suppose S and S' are sentences of natural language, that S' serves as an analysis of S, and that S and S' exhibit hyperintensionality (i.e., there are contexts of natural language in which S and S' are not substitutable for one another, despite the fact that analysis shows them to be necessarily equivalent). Then if formulas φ and ψ are formal representations of S and S', respectively, within our system, we might capture that analysis by way of a definition of the form $\varphi \equiv_{df} \psi$. But we wouldn't want to capture the analysis in terms of the definition of the form $\varphi =_{df} \psi$, since that would make it impossible to represent the contexts of natural language in which S and S' exhibit hyperintensionality. So, unless there is a good reason to suppose that the definiens and definiendum in a formula definition bear a stronger connection than necessary equivalence, we'll continue to introduce definitions-by-equivalence. Such definitions yield only the modal equivalence $\Box(\varphi \equiv \psi)$ from the definition $\varphi \equiv_{df} \psi$, and not an identity.[147]

(**110**) **Metarules/Derived Rules:** Rule for the Substitution of Identicals (Rule =E). We now have:

> **Rule =E**
> If $\Gamma_1 \vdash \varphi^\tau_\alpha$ and $\Gamma_2 \vdash \tau = \sigma$, then $\Gamma_1, \Gamma_2 \vdash \varphi'$, whenever τ and σ are any terms substitutable for α in φ, and φ' is the result of replacing zero or more occurrences of τ in φ^τ_α with occurrences of σ. [Variant: $\varphi^\tau_\alpha,\ \tau = \sigma \vdash \varphi'$]

In the usual manner, we may conceive of the Variant of Rule =E as a derived rule that lets us infer formulas from formulas. Note also that Rule =E governs

[146] This observation applies to the present, second-order version of object theory. In typed object theory, not every formula φ is significant, and so one may not derive $\varphi = \psi$ from any definition-by-identity $\varphi =_{df} \psi$; one can only derive such identities when ψ is significant. Thus, in typed object theory, whereas every definition by equivalence yields a well-formed biconditional theorem, not every definition by identity of a 0-ary relation term yields an identity as a theorem and so we can't always derive an equivalence from a definition by identity. Typed object theory won't have the option of eliminating definitions by equivalence in favor of definitions by identity.

[147] If one exercises the option of using $=_{df}$ to define new formulas, then it should be remembered that the following is *not* a valid inference: $\varphi = (\psi\ \&\ \chi),\ \psi \vdash \varphi = \chi$. In general, there is no rule of simplification for identity that corresponds to the Rule $\equiv$S of Biconditional Simplification (91).

any terms τ and σ. Since $\tau = \sigma$ is an assumption of the rule, the rule is applicable only when both terms are significant, given theorems (107.1) and (107.2). Note finally that as a special case of the rule, we have $\varphi^\tau_\alpha, \tau = \sigma \vdash \varphi^\sigma_\alpha$. This is the case where φ' is φ^σ_α, i.e., the result of replacing *all* of the occurrences of τ in φ^τ_α by σ.

(**111**) **Theorems:** Special Cases of Conversion for 0-ary λ-Expressions and Other Facts. Theorem (104.2) allows us to prove, as *theorem schemata*, the following unconditional conversion principles governing 0-ary λ-expressions and formulas: (.1) *that*-φ is identical to φ; (.2) *that*-φ is true if and only if φ; (.3) *that*-φ and its alphabetic variants are identical; (.4) φ and its alphabetic variants are identical; (.5) φ and its alphabetic variants are equivalent; and (.6) φ is equivalent to ψ if and only if *that*-φ is equivalent to *that*-ψ. Formally:

(.1) $[\lambda\, \varphi] = \varphi$, for any formula φ (special case of η-Conversion)

(.2) $[\lambda\, \varphi] \equiv \varphi$, for any formula φ (special case of β-Conversion)

(.3) $[\lambda\, \varphi] = [\lambda\, \varphi]'$, where $[\lambda\, \varphi]'$ is any alphabetic variant of $[\lambda\, \varphi]$ (special case of α-Conversion)

(.4) $\varphi = \varphi'$, where φ' is any alphabetic variant of φ

(.5) $\varphi \equiv \varphi'$, where φ' is any alphabetic variant of φ

(.6) $(\varphi \equiv \psi) \equiv ([\lambda\, \varphi] \equiv [\lambda\, \psi])$

(.3) is a simple consequence of axiom (39.2) and the 0-ary case of the axiom α-Conversion (48.1). (.4) is derivable from (.3) and (.1).

(**112**) **Remark:** The Theory of Truth. It should be observed that (111.2) constitutes a theory of truth. In this theorem, the λ expresses 'that' (since it binds no variables) and the expression $[\lambda\, \varphi]$ must be read as a formula (since it stands to the left of the biconditional sign). We observed, in (3), that since *truth* is the 0-ary case of predication, we should read (111.2) as: *that*-φ is true if and only if φ. Assuming we've applied our theory in the usual way, then the instance $[\lambda\, Po] \equiv Po$ might assert *that-Obama-is-President* is true if and only if Obama is President. With (111.2), we have established that the propositional version of the Tarski T-Schema is a *theorem*![148] Some philosophers may regard this as a *deflationary* theory of truth, since the concept of *truth* is, in some sense, eliminated; the notion of truth is expressed one one side of (111.2) but not on the other. Other philosophers may regard this as a *primitivist* theory of truth, since *truth* has been reduced to the primitive, 0-ary case of exemplification

[148] It is also a tautology; see Zalta 2014, where this was first noted, as well as Remark (92) and the next Remark.

predication. Indeed, since $\varphi\downarrow$ (104.2) and $[\lambda\, \varphi] \equiv \varphi$ (111.2) hold even when φ is an encoding formula, we might say that truth is just the limiting case of predication.

(**113**) **Remark:** On the Derivability of Tautologies. The derivation of (111.1) and (111.2) is needed for the proof of Metatheorem ⟨9.1⟩, which asserts that every valuation of the *prime* formulas (i.e., formulas *not* having the form $\neg\psi$, $\psi \to \chi$, or $[\lambda\, \psi]$) induces a certain derivability relation. This metatheorem is the key lemma needed for the proof of Metatheorem ⟨9.2⟩, i.e., that every tautology is derivable. See the proofs of these metatheorems in the Appendix to this chapter.

(**114**) **Metarule/Derived Rule:** Rule of Alphabetic Variants. The classical Rule of Alphabetic Variants is now easily established from the recently proved theorem (111.5). The classical rule asserts that φ is derivable from Γ if and only if any of its alphabetic variants is derivable from Γ:

> **Rule of Alphabetic Variants**
> $\Gamma \vdash \varphi$ if and only if $\Gamma \vdash \varphi'$, where φ' is any alphabetic variant of φ
> [Variant $\varphi \dashv\vdash \varphi'$]

As a special case, when $\Gamma = \varnothing$, our rule asserts that a formula is a theorem if and only if all of its alphabetic variants are theorems. We henceforth use the Variant form within derivations as a derived rule.

Now that it is established that every pair of alphabetically-variant formulas have the same inferential role, we no longer need to adopt the convention that bound object-language variables in definientia function as metavariables; i.e., convention (17.2.b) is no longer strictly necessary. If we use object-language variables for the bound variables in a definiens, then the alphabetic variants of any theorem introduced by the resulting definition will also be a theorem. This fact was mentioned in Remark (28), where we discussed why bound variables in definientia should be, or function as, metavariables.

(**115**) **Theorems:** Individuals Are Ordinary or Abstract. We now prove that (.1) *being ordinary* exists; (.2) *being abstract* exists; (.3) x is ordinary if and only if it is possible that x is concrete; (.4) x is abstract if and only if x couldn't be concrete; and (.5) an object is either ordinary or abstract:

(.1) $O!\downarrow$

(.2) $A!\downarrow$

(.3) $O!x \equiv \Diamond E!x$

(.4) $A!x \equiv \neg\Diamond E!x$

(.5) $O!x \vee A!x$

We observe here that (.5) is one of the elements needed to prove that identity is a reflexive condition on individuals, which forms part of the proof that identity is generally reflexive (117.1).

(**116**) **Theorems:** Facts About Relation Identity. Our definitions for relation identity and theory of definitions yield that (.1) F^1 is identical to G^1 if and only if necessarily, all and only the individuals that encode F^1 encode G^1; (.2) F^n is identical to G^n if and only if each way of projecting F^n and G^n onto $n-1$ objects results in identical properties; and (.3) p is identical to q if and only if *being such that* p is identical to *being such that* q:

(.1) $F^1 = G^1 \equiv \Box\forall x(xF^1 \equiv xG^1)$ $(n \geq 1)$

(.2) $F^n = G^n \equiv$ $(n \geq 2)$

$$\forall y_1 \ldots \forall y_{n-1}([\lambda x\, F^n x y_1 \ldots y_{n-1}] = [\lambda x\, G^n x y_1 \ldots y_{n-1}]\ \& \\ [\lambda x\, F^n y_1 x y_2 \ldots y_{n-1}] = [\lambda x\, G^n y_1 x y_2 \ldots y_{n-1}]\ \& \ldots \& \\ [\lambda x\, F^n y_1 \ldots y_{n-1} x] = [\lambda x\, G^n y_1 \ldots y_{n-1} x])$$

(.3) $p = q \equiv [\lambda x\, p] = [\lambda x\, q]$

These equivalences were used as definitions in previous publications on object theory, where less attention was paid to the issues surrounding non-denoting complex terms.

It is important not to misconstrue these theorems as principles that *determine* whether arbitrary relations F and G are identical. Some relation identities have to discovered *a posteriori*, and no *a priori* principle can be used as a substitute. For example, *being a woodchuck* = *being a groundhog* is an identity claim about properties that is discovered *a posteriori* and not determined by an *a priori* principle, just as Hesperus = Phosphorus is an identity claim about individuals that is discovered *a posteriori* rather than determined by an *a priori* principle. So, in an important sense, these theorems are *not* 'criteria of identity' for relations. Rather, what they do is tell us (a) what it is we know when we assert or prove that $F = G$, (b) what we have to prove if we are to show that relations (represented in our system) are identical, and (c) what the consequences are of asserting in our system that $F = G$ (e.g., what the consequences are of adding *being a woodchuck* = *being a groundhog* as an axiom or hypothesis to our system). So these are principles that govern relations by placing constraints on their inferential role within PLM. For further discussion, see Remark (190).

(**117**) **Theorems:** Identity is an Equivalence Condition. Our system yields, as theorems, that identity is an equivalence condition on individuals and on relations of every arity. We can express this using schemata, in which α, β, and γ are any individual variable or any n-ary relation variables for some n:

(.1) $\alpha = \alpha$

(.2) $\alpha = \beta \rightarrow \beta = \alpha$

(.3) $\alpha = \beta \;\&\; \beta = \gamma \rightarrow \alpha = \gamma$

Clearly, it also follows that:

(.4) $\alpha = \beta \equiv \forall\gamma(\alpha = \gamma \equiv \beta = \gamma)$

Exercises: Prove that (.2) holds for any two terms τ and τ' having the same type, and prove that (.3) holds for any three terms, τ, τ', and τ'', all of the same type. Explain why (.4) fails in the right-to-left direction when arbitrary terms of the same type τ and σ replace α and β, respectively, in the statement of the claim, i.e., show that $\tau = \sigma \equiv \forall\alpha(\tau = \alpha \equiv \sigma = \alpha)$ fails in the right-to-left direction. Finally, show that $\tau\downarrow \rightarrow (\tau = \sigma \equiv \forall\alpha(\tau = \alpha \equiv \sigma = \alpha))$.

(**118**) **Metarules/Derived Rules:** Rules of Identity Introduction (Rule =I). The following rules govern the terms of our language:

(.1) **Rule =I**
If $\Gamma \vdash \tau\downarrow$, then $\Gamma \vdash \tau = \tau$ [Variant: $\tau\downarrow \vdash \tau = \tau$]

(.2) **Rule =I** (Special Case)
$\vdash \tau = \tau$, provided τ is a primitive constant, a variable, or a core λ-expression.

The Variant of (.1) tells us that whenever we take $\tau\downarrow$ as an assumption when reasoning, we may infer $\tau = \tau$.[149]

It is worth observing here that from the Variant rule $\tau\downarrow \vdash \tau = \tau$, it follows by the Deduction Theorem (75) that $\tau\downarrow \rightarrow \tau = \tau$ is a theorem. And since we know, by an instance of (107.1), that $\tau = \tau \rightarrow \tau\downarrow$, it follows that $\tau\downarrow \equiv \tau = \tau$. So for any definite description, it follows that $\imath x\varphi\downarrow \equiv (\imath x\varphi = \imath x\varphi)$. (Recall our convention, mentioned at the end of (20), that $\imath x\varphi\downarrow$ abbreviates $(\imath x\varphi)\downarrow$.) This captures principle $*14\cdot28$ in Whitehead and Russell 1910–1913 ([1925–1927], pp. 175, 184):

$$*14\cdot 28.\vdash : E!(\imath x)(\varphi x)\,.\equiv.\,(\imath x)(\varphi x) = (\imath x)(\varphi x)$$

It is interesting that Whitehead and Russell read this claim as: "$(\imath x)(\varphi x)$ only satisfies the reflexive property of identity if $(\imath x)(\varphi x)$ exists" (175). Given their discussion on pp. 30–31 ([1925–1927], Chapter I), it should be clear that our formula $\imath x\varphi\downarrow$ is the counterpart to Whitehead & Russell's formula $E!(\imath x)(\varphi x)$.

[149] To see that Rule =I and its Variant are equivalent, we derive the Variant from the Rule and vice versa. ($\hookrightarrow$) Assume the Rule, i.e., assume that if $\Gamma \vdash \tau\downarrow$, then $\Gamma \vdash \tau = \tau$. Consider the instances in which Γ consists of the sole premise $\tau\downarrow$: if $\tau\downarrow \vdash \tau\downarrow$, then $\tau\downarrow \vdash \tau = \tau$. But we know the antecedent of this last claim, by the special case of (63.2). Hence, $\tau\downarrow \vdash \tau = \tau$. ($\hookleftarrow$) Assume the Variant: $\tau\downarrow \vdash \tau = \tau$. Then $\vdash \tau\downarrow \rightarrow \tau = \tau$, by the Deduction Theorem/Conditional Proof (75). Now to see that Rule =I holds, assume its antecedent: $\Gamma \vdash \tau\downarrow$. It then follows from what we've established and assumed that $\Gamma \vdash \tau = \tau$, by (63.5).

(**119**) **Remark:** Preparing the Ground for the Rule of Identity by Definition. We now show how to derive the classical inferential role of definitions-by-= for the case in which no variable occurs free in the definiens and definiendum. It should then be clear how this generalizes to the case in which there are free variables in the definition and so prepare us for the general statement of the metarule in (120).

When there are no free variables in the definiens or definiendum, a definition-by-= has the simple form $\tau =_{df} \sigma$ and introduces τ as a new constant of the same type as σ. We saw in (73) that in this case, the primitive Rule of Definition by Identity implies that the following is an axiom:

$$(\sigma\downarrow \rightarrow \tau = \sigma) \;\&\; (\neg\sigma\downarrow \rightarrow \neg\tau\downarrow)$$

So whenever $\sigma\downarrow$ is a theorem or a hypothesis it follows that $\tau = \sigma$. Thus, the inferential role of a definition-by-= without free variables reduces to the classical one; if it is provable, by modally strict means, that the definiens is significant, then the definition $\tau =_{df} \sigma$ yields $\tau = \sigma$ as a $\square$-theorem and, moreover, yields that $\tau = \sigma$ follows from any set of premises, by (63.3). This reasoning now generalizes to the following derived metarule.

(**120**) **Metarule:** The Rule of Identity by Definition. Let us again use $\sigma(\tau_1, \ldots, \tau_n)$ and $\tau(\tau_1, \ldots, \tau_n)$ to abbreviate $\sigma^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$ and $\tau^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$, respectively $(1 \leq i \leq n)$. Then the following derived metarule is justified:

(.1) **Rule of Identity by Definition**
Whenever $\tau(\alpha_1, \ldots, \alpha_n) =_{df} \sigma(\alpha_1, \ldots, \alpha_n)$ is a definition-by-= in which $\alpha_1, \ldots, \alpha_n$ occur free $(n \geq 0)$ and $\tau_1, \ldots, \tau_n$ are substitutable for $\alpha_1, \ldots, \alpha_n$, respectively, in both definiens and definiendum, then:

$$\text{if } \Gamma \vdash \sigma(\tau_1, \ldots, \tau_n)\downarrow, \text{ then } \Gamma \vdash \tau(\tau_1, \ldots, \tau_n) = \sigma(\tau_1, \ldots, \tau_n)$$

By convention, we omit the identical form of the rule for $\vdash_\square$. The justification of this metarule, which is straightforward, is in the Appendix.

It immediately follows that, for any instance of a definition-by-=, when the definiens is significant, we have the classical introduction and elimination rules for the definiendum:

(.2.a) **Rule of Definiendum Elimination** (Rule $=_{df}$**E**):
Whenever $\tau(\alpha_1, \ldots, \alpha_n) =_{df} \sigma(\alpha_1, \ldots, \alpha_n)$ is a definition-by-= in which $\alpha_1, \ldots, \alpha_n$ occur free $(n \geq 0)$, $\tau_1, \ldots, \tau_n$ are substitutable for $\alpha_1, \ldots, \alpha_n$, respectively, in both the definiens and definiendum, φ contains one or more occurrences of $\tau(\tau_1, \ldots, \tau_n)$, and φ' is the result of replacing zero or more occurrences of $\tau(\tau_1, \ldots, \tau_n)$ in φ by $\sigma(\tau_1, \ldots, \tau_n)$, then if $\Gamma \vdash \sigma(\tau_1, \ldots, \tau_n)\downarrow$ and $\Gamma \vdash \varphi$, then $\Gamma \vdash \varphi'$.

[Variant: $\sigma(\tau_1, \ldots, \tau_n)\downarrow, \varphi \vdash \varphi'$]

(.2.b) **Rule of Definiendum Introduction** (Rule $=_{df}$**I**):
Whenever $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$ is a definition-by-= in which α_1, $\ldots,\alpha_n$ occur free ($n \geq 0$), $\tau_1,\ldots,\tau_n$ are substitutable for $\alpha_1,\ldots,\alpha_n$, respectively, in both definiens and definiendum, φ contains one or more occurrences of $\sigma(\tau_1,\ldots,\tau_n)$, and φ' is the result of replacing zero or more occurrences of $\sigma(\tau_1,\ldots,\tau_n)$ in φ by $\tau(\tau_1,\ldots,\tau_n)$, then if $\Gamma \vdash \sigma(\tau_1,\ldots,\tau_n)\downarrow$ and $\Gamma \vdash \varphi$, then $\Gamma \vdash \varphi'$.

[Variant: $\sigma(\tau_1,\ldots,\tau_n)\downarrow, \varphi \vdash \varphi'$]

And similarly for $\vdash_\Box$.

Once we gain some experience citing these rules when proving theorems, we'll then revert to the classical citation 'by definition' when reasoning from definitions-by-=. Clearly, if $\tau(\tau_1,\ldots,\tau_n) =_{df} \sigma(\tau_1,\ldots,\tau_n)$ is an instance of a definition-by-=, then reasoning is classical when $\sigma(\tau_1,\ldots,\tau_n)$ is significant; we can substitute $\tau(\tau_1,\ldots,\tau_n)$ and $\sigma(\tau_1,\ldots,\tau_n)$ for one another in *any* context.

(**121**) **Theorems:** The Definitions and Axioms of Negative Free Logic Derived as Theorems.

(.1) $\tau\downarrow \equiv \exists\beta(\beta = \tau)$, provided that β doesn't occur free in τ

(.2) $\forall\alpha\varphi \rightarrow (\exists\beta(\beta = \tau) \rightarrow \varphi^\tau_\alpha)$, provided τ is substitutable for α in φ and β doesn't occur free in τ

(.3) $\exists\beta(\beta = \tau)$, provided (a) τ is either a primitive constant, a variable, or a core λ-expression, and (b) β doesn't occur free in τ

(.4) $(\Pi^n\kappa_1\ldots\kappa_n \vee \kappa_1\ldots\kappa_n\Pi^n) \rightarrow \exists\beta(\beta = \tau)$, where τ is any of Π^n, κ_1, ..., or κ_n, and β doesn't occur free in τ.

(**122**) **Remark:** Digression on the Proviso for (121). Each of the theorems in (121) require the proviso that β doesn't occur free in τ. To see why it is required for the left-to-right direction of (121.1), consider the fact that the proviso rules out $[\lambda x\, \neg Fx]\downarrow \rightarrow \exists F(F = [\lambda x\, \neg Fx])$ as an instance, since the variable F ($= \beta$) occurs free in $[\lambda x\, \neg Fx]$ ($= \tau$). Call this formula (ϑ). If (ϑ) had been a theorem, we would have been able to derive a contradiction, as follows. By (39.2), we know $[\lambda x\, \neg Fx]\downarrow$, since this is a core λ-expression. Hence, it would follow from this and (ϑ) that $\exists F(F = [\lambda x\, \neg Fx])$. Suppose P were a witness to $\exists F(F = [\lambda x\, \neg Fx])$, so that $P = [\lambda x\, \neg Px]$ and, by symmetry, $[\lambda x\, \neg Px] = P$. Now since $[\lambda x\, \neg Px]$ is also a core λ-expression, (39.2) would also yield $[\lambda x\, \neg Px]\downarrow$. So by an appropriate instance of β-Conversion (48.2), it would follow that:

(∗) $[\lambda x\, \neg Px]y \equiv \neg Py$

But since $[\lambda x\, \neg Px] = P$, Rule =E would allow us to substitute P for $[\lambda x\, \neg Px]$ in (∗) to obtain $Py \equiv \neg Py$, which would be a contradiction.

The reason why the proviso that β doesn't occur free in τ is required for the right-to-left direction of (121.1) is given in the proof of this direction in the Appendix; in particular, see footnote 432.

Given our explanation as to why the proviso is needed for the left-to-right direction of (121.1), it should be clear why the proviso governs (121.3); we can't allow $\exists F(F = [\lambda x\, \neg Fx])$ to be an instance of (121.3), by the reasoning just described. Intuitively, such a formula asserts the falsehood that there is a property that is identical to its own negation.

As to (121.2) and (121.4), we leave it to the reader to come up with an explanation as to why the proviso is needed. But note that in (121.2), if β were to have a free occurrence in τ, then if we were to apply (121.2) by using $\forall\alpha\varphi$ and $\exists\beta(\beta = \tau)$ to conclude φ^{τ}_{α}, the latter formula would have a free occurrence of β. While φ^{τ}_{α} (with a free occurrence of β) might hold for some values of β, it is not guaranteed to hold universally for all values of β. But, if we were to derive such a formula as a theorem, Rule GEN would let us conclude $\forall\beta\varphi^{\tau}_{\alpha}$.

(**123**) **Theorems:** (Necessarily) Every Individual/Relation (Necessarily) Exists. Where α and β are distinct variables of the same type, it is a consequence of our axioms and rules that:

(.1) (.a) $\forall\alpha\, \alpha\!\downarrow$
(.b) $\forall\alpha\exists\beta(\beta = \alpha)$

(.2) (.a) $\Box\alpha\!\downarrow$
(.b) $\Box\exists\beta(\beta = \alpha)$

(.3) (.a) $\Box\forall\alpha\, \alpha\!\downarrow$
(.b) $\Box\forall\alpha\exists\beta(\beta = \alpha)$

(.4) (.a) $\forall\alpha\Box\alpha\!\downarrow$
(.b) $\forall\alpha\Box\exists\beta(\beta = \alpha)$

(.5) (.a) $\Box\forall\alpha\Box\alpha\!\downarrow$
(.b) $\Box\forall\alpha\Box\exists\beta(\beta = \alpha)$

Note that when α and β are the individual variables x and y, we have the following instances of (.1.a) and (.1.b), respectively:

$\forall x\, x\!\downarrow$
$\forall x\exists y(y = x)$

The first asserts that every individual exists, while the second asserts that every individual is such that there exists something identical to it. Similarly, the following are instances of (.2.a) and (.2.b), respectively:

$\Box x{\downarrow}$
$\Box\exists y(y=x)$

The first asserts that necessarily x exists, while the second asserts that necessarily there exists an individual identical to x. These are distinct modal claims. Strictly speaking, we should not read $\exists y(y=x)$ as *x exists*, for otherwise we would collapse the readings of the (.a) and (.b) forms of the above theorems. $x{\downarrow}$ and $\exists y(y=x)$ make distinct claims. Since they are equivalent (121.1), it often does no harm, in some contexts, to slide between them. But the former is literally defined as an *existence* claim (20.1) while the latter is literally defined as an *existentially quantified* identity claim (23.1).[150]

(**124**) **Theorems:** Self-Identity and Necessity. It is a consequence of the foregoing that (.1) necessarily everything is self-identical, and that (.2) everything is necessarily self-identical. Where α is any variable:

(.1) $\Box\forall\alpha(\alpha=\alpha)$

(.2) $\forall\alpha\Box(\alpha=\alpha)$

These well-known principles of self-identity and necessity are thus provable.

(**125**) **Theorems:** Necessity of Identity. Where α and β are any variables of the same type, and τ and σ are any terms of the same type, it is provable that (.1) if α and β are identical, then necessarily they are identical, and (.2) if τ and σ are identical, then necessarily they are identical:

(.1) $\alpha=\beta \rightarrow \Box\alpha=\beta$, provided α and β are variables of the same type

(.2) $\tau=\sigma \rightarrow \Box\tau=\sigma$, provided τ and σ are terms of the same type

When α and β are the individual variables x and y, (.1) asserts $x=y \rightarrow \Box x=y$. This is the famous principle of the *necessity of identity* for individuals, discussed at length in Kripke 1971. But (.1) has greater significance than the principle that Kripke discusses: (a) it is derived in a more general form as a theorem schema whose instances apply to any individual or relation, and (b) it doesn't assume identity as a primitive. We've already seen that the *definitions* of object identity in (23.1) and relation identity in (23.2), (23.3), and (23.4) ground the reflexivity of identity (117.1). The reflexivity of identity is one of the key facts used in the proof of (.1).

(**126**) **Theorems:** Quantification and Identity.

[150]Thus, given the present work, one should use the label Necessary Existence (NE) to designate the instance $\forall x\Box x{\downarrow}$ of (.4.a) ('everything necessarily exists') and *not* the instance $\forall x\Box\exists y(y=x)$ of (.4.b). This differs from Linsky & Zalta 1994 (435), where identity was taken as a primitive and the label (NE) was used to designate $\forall x\Box\exists y(y=x)$. Though the present work preserves the fixed-domain understanding of quantified S5 modal logic defended in Linsky & Zalta 1994, we would now read its theorems a bit differently.

(.1) $\varphi \equiv \exists\beta(\beta = \alpha \;\&\; \varphi^{\beta}_{\alpha})$, provided β is substitutable for α in φ and doesn't occur free in φ.

(.2) $\tau\!\downarrow \rightarrow (\varphi^{\tau}_{\alpha} \equiv \exists\alpha(\alpha = \tau \;\&\; \varphi))$, provided τ is substitutable for α in φ

(.3) $(\varphi \;\&\; \forall\beta(\varphi^{\beta}_{\alpha} \rightarrow \beta = \alpha)) \equiv \forall\beta(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha)$,
provided α, β are distinct variables of the same type, and β is substitutable for α in φ and doesn't occur free in φ.

(.4) $(\varphi^{\beta}_{\alpha} \;\&\; \forall\alpha(\varphi \rightarrow \alpha = \beta)) \equiv \forall\alpha(\varphi \equiv \alpha = \beta)$,
provided α, β are distinct variables of the same type, and β is substitutable for α in φ and doesn't occur free in φ.

A simple example of (.1) is $Px \equiv \exists y(y = x \;\&\; Py)$, and a simple example of (.2) is $a\!\downarrow \rightarrow (Qa \equiv \exists x(x = a \;\&\; Qx))$. But these theorems also apply to relation terms; as examples we have $Fa \equiv \exists G(G = F \;\&\; Ga)$ and $P\!\downarrow \rightarrow (Px \equiv \exists F(F = P \;\&\; Fx))$, respectively. The reader should construct examples in which φ has greater complexity. Note that the antecedent of (.2) restricts the consequent to significant terms. To see why, let φ be $\neg Px$, α be x, and τ be $\imath yQy$, so that φ^{τ}_{α} is $\neg P\imath yQy$. Then if $\neg(\imath yQy)\!\downarrow$, it is easy to show that $\neg P\imath yQy$ is provably true but $\exists x(x = \imath yQy \;\&\; \neg Px)$ is provably false (exercise).

Theorem (.3) and (.4) are noteworthy; in each case, the two sides of the main biconditionals are equivalent ways of asserting an important claim. Consider (.3). When α is an individual variable, both sides of the biconditional are true whenever *α is a unique individual such that φ*, and when α is a relation variable, both sides are true whenever *α is a unique relation such that φ*. (In the formal mode, we would say that both sides of the biconditional are true if and only if α uniquely satisfies φ.) Similar remarks apply to (.4), except that both sides of the biconditional intuitively assert that β uniquely satisfies φ.

(**127**) **Definition and Theorem:** Uniqueness Quantifier. In light of the last observation, there are two equivalent ways to define the special *uniqueness* quantifier '$\exists!$' so that formulas of the form $\exists!\alpha\varphi$ assert that there exists a unique α such that φ. We officially define this notion as (.1) some α is such that φ and every entity such that φ is identical to α. And then we prove, as a theorem, that (.2) there exists an α such that all and only the entities such that φ are identical to α:

(.1) $\exists!\alpha\varphi \;\equiv_{df}\; \exists\alpha(\varphi \;\&\; \forall\beta(\varphi^{\beta}_{\alpha} \rightarrow \beta = \alpha))$,
provided β doesn't occur free, and is substitutable for α, in φ

(.2) $\exists!\alpha\varphi \equiv \exists\alpha\forall\beta(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha)$,
provided β doesn't occur free, and is substitutable for α, in φ

The readers should take care not to confuse the defined, variable-binding, formula-forming operator '$\exists!$' with the primitive, unary relation constant '$E!$'.

(**128**) **Theorems:** Uniqueness Implies At Most One. It follows from the definition of the uniqueness quantifier that if there exists a unique φ, then at most one entity is such that φ:

$$\exists!\alpha\varphi \rightarrow \forall\beta\forall\gamma((\varphi^{\beta}_{\alpha} \,\&\, \varphi^{\gamma}_{\alpha}) \rightarrow \beta = \gamma),$$
provided β and γ don't occur free, and are substitutable for α, in φ

In this theorem schema, α, β, and γ are either all individual variables or all n-ary relation variables, for some n.

(**129**) **Theorems:** Uniqueness and Necessity. It is now provable that if every entity such that φ is necessarily such that φ, then if there is a unique entity such that φ, there is a unique entity necessarily such that φ:

$$\forall\alpha(\varphi \rightarrow \Box\varphi) \rightarrow (\exists!\alpha\varphi \rightarrow \exists!\alpha\Box\varphi)$$

In other words, if for every α, φ *necessarily holds* of α whenever it holds of α, then if there is exactly one entity such that φ, there is exactly one entity necessarily such that φ.

9.8 The Theory of Actuality and Descriptions

Although the theorems in this section sometimes involve the necessity operator, no special principles for necessity other than the axioms and rules introduced thus far are required to prove the basic theorems and metarules governing actuality and descriptions.

9.8.1 The Theory of Actuality

We first prove two ⋆-theorems that are derivable from the modally fragile axiom $\mathcal{A}\varphi \rightarrow \varphi$ (43)⋆, and then focus to several groups of modally strict theorems that are needed to prove the Rule of Actualization. Finally, we'll develop a variety of other theorems about actuality, most of which are modally strict.

(**130**) ⋆**Theorems:** It follows from axiom (43)⋆ that (.1) if φ, then it is actually the case that φ, and that (.2) it is actually the case that φ if and only if φ:

(.1) $\varphi \rightarrow \mathcal{A}\varphi$

(.2) $\mathcal{A}\varphi \equiv \varphi$

It is relatively straightforward to develop an intuitive, semantic argument as to why (.1) can't be necessitated, i.e., to describe models in which the necessitation of (.1) fails to be true.[151]

[151] Consider models of the theory in which there are two possible worlds, $\boldsymbol{w}_{\alpha}$ and $\boldsymbol{w}_1$, and where φ is false at the actual world $\boldsymbol{w}_{\alpha}$ but true at $\boldsymbol{w}_1$. Then $\Box(\varphi \rightarrow \mathcal{A}\varphi)$ is false at $\boldsymbol{w}_{\alpha}$ because $\varphi \rightarrow \mathcal{A}\varphi$ is false at $\boldsymbol{w}_1$ — at $\boldsymbol{w}_1$, φ is true and $\mathcal{A}\varphi$ is false.

(**131**) **Theorems:** Actuality Distributes Over a Conditional. It follows from axiom (44.2) that if it is actually the case that if-φ-then-ψ, then if it is actually the case that φ, then it is actually the case that ψ:

$$\mathcal{A}(\varphi \to \psi) \to (\mathcal{A}\varphi \to \mathcal{A}\psi)$$

This is analogous to the K axiom for the modal operator $\Box$; the actuality operator distributes over a conditional.

(**132**) **Theorems:** Necessity Implies Actuality. It is straightforward to show that if necessarily φ, then actually φ:

$$\Box\varphi \to \mathcal{A}\varphi$$

Note that in the Appendix, we do *not* give the following proof of this theorem: from the assumption $\Box\varphi$, infer φ by the T schema; then infer $\mathcal{A}\varphi$ by the modally fragile theorem (130.1)★; then and conclude $\Box\varphi \to \mathcal{A}\varphi$ by conditional proof. Though this is a perfectly good proof, it is not modally strict. By contrast, the proof in the Appendix is modally strict and so one may apply RN to our theorem to obtain $\Box(\Box\varphi \to \mathcal{A}\varphi)$.

(**133**) **Theorems:** Actuality, Conjunctions, and Biconditionals. The following theorems also have modally strict proofs. (.1) it is actually the case that if actually φ then φ; (.2) it is actually the case that if φ then actually φ; (.3) if is actually the case that φ and actually the case that ψ, then it is actually the case that both φ and ψ; and (.4) it is actually the case that, actually φ if and only if φ:

(.1) $\mathcal{A}(\mathcal{A}\varphi \to \varphi)$

(.2) $\mathcal{A}(\varphi \to \mathcal{A}\varphi)$

(.3) $(\mathcal{A}\varphi \,\&\, \mathcal{A}\psi) \to \mathcal{A}(\varphi \,\&\, \psi)$

(.4) $\mathcal{A}(\mathcal{A}\varphi \equiv \varphi)$

(.1) tells us that the actualizations of instances of the modally fragile axiom (43)★ are modally strict theorems. We should note that the right-to-left direction of (.3) will be established later, as the left-to-right direction of (139.2).

(**134**) **Theorems:** Actualizations and Universal Closures of the Previous Theorem. Note that by the left-to-right direction of axiom (44.4), $\mathcal{A}\varphi \to \mathcal{A}\mathcal{A}\varphi$. If we apply this principle to theorem (133.1) and repeat the process, then (.1) the closures of $\mathcal{A}\varphi \to \varphi$ obtained by prefacing any string of $\mathcal{A}$ operators is derivable:

(.1) $\mathcal{A}\ldots\mathcal{A}(\mathcal{A}\varphi \to \varphi)$, for any finite string of actuality operators $\mathcal{A}\ldots\mathcal{A}$

Moreover, by applying GEN to (133.1), we know (.2) every α is such that it is actually the case that: if actually φ then φ:

(.2) $\forall\alpha\mathcal{A}(\mathcal{A}\varphi \to \varphi)$

Now it follows from (.2) by axiom (44.3) that (.3) actually, every α is such that: if actually φ then φ:

(.3) $\mathcal{A}\forall\alpha(\mathcal{A}\varphi \to \varphi)$

Hence the actualization of a universal closure of an instance of axiom (43)★ is derivable by modally strict means. Moreover, by repeating the steps in the proofs of (.2) and (.3) enough times, it follows that:

(.4) $\mathcal{A}\forall\alpha_1 \dots \forall\alpha_n(\mathcal{A}\varphi \to \varphi)$

By (.4), the actualization of any universal closure of the axiom schema $\mathcal{A}\varphi \to \varphi$ (43)★ is a modally strict theorem. Thus, given (133.1) and (.4), the actualizations of all the axioms asserted in (43)★ are modally strict theorems. This fact is needed in the proof of the Rule of Actualization, to which we now turn.

Exercise: Show that (.1) – (.4) all hold with respect to $\mathcal{A}\varphi \equiv \varphi$, i.e., show that:

- $\mathcal{A}\dots\mathcal{A}(\mathcal{A}\varphi \equiv \varphi)$, for any finite string of actuality operators $\mathcal{A}\dots\mathcal{A}$
- $\forall\alpha\mathcal{A}(\mathcal{A}\varphi \equiv \varphi)$
- $\mathcal{A}\forall\alpha(\mathcal{A}\varphi \equiv \varphi)$
- $\mathcal{A}\forall\alpha_1 \dots \forall\alpha_n(\mathcal{A}\varphi \equiv \varphi)$

(**135**) **Metadefinition and Metarule:** Rule of Actualization (RA). We first define:

- $\mathcal{A}\Gamma = \{\mathcal{A}\psi \mid \psi \in \Gamma\}$ (Γ any set of formulas)

Thus, $\mathcal{A}\Gamma$ is the result of prefixing the actuality operator to every formula in Γ. We then have the following metarule:

Rule of Actualization (RA)
If $\Gamma \vdash \varphi$, then $\mathcal{A}\Gamma \vdash \mathcal{A}\varphi$

We most often use this rule in the form in which Γ is empty:

- If $\vdash \varphi$, then $\vdash \mathcal{A}\varphi$

In other words, whenever φ is a theorem, so is $\mathcal{A}\varphi$.

By the convention in Remark (67), we omit the $\vdash_\Box$ version of RA. However, the justification of RA in the Appendix makes it clear that the $\vdash_\Box$ version of the rule is easier to justify.[152]

[152] The base case of the inductive justification of RA (in the Appendix) considers one-step derivations of φ from Γ in which φ is an axiom. So it has to consider the case where φ is an instance of (43)★. But for the $\vdash_\Box$ version of RA, the base step would consider the *modally strict* one-step derivations of φ from Γ in which φ is an axiom. In that case, φ can't be an instance of (43)★; there can't be a modally strict one-step derivation of φ from Γ in which φ is an instance of (43)★.

Note also that should one wish to extend our system with new axioms, RA can easily be preserved as a justified metarule as long as we either (a) axiomatically assert the actualizations of the new axioms or (b) show, as in the case of axiom (43)★, that the actualizations of the new axioms are provable as theorems.[153]

(**136**) **Remark:** Digression on the Formulation of RA. It is important to recognize why RA is formulated as in (135), as opposed to the following alternative:

If $\Gamma \vdash \varphi$, then $\Gamma \vdash \mathcal{A}\varphi$

The consequent of this rule differs from the consequent of RA by stating that there is a derivation of $\mathcal{A}\varphi$ from Γ rather than from $\mathcal{A}\Gamma$. One can prove that this version of the rule is semantically valid.[154] However, the justification of this rule depends on the modally fragile axiom of actuality (43)★, i.e., (43)★ is *used* in the justification.[155] Clearly, the application of this alternative rule would undermine modally-strict derivations. For example, given that $\varphi \vdash \varphi$ is an instance of metarule (63.2), this alternative version of RA would allow us to conclude $\varphi \vdash \mathcal{A}\varphi$, which by the Deduction Theorem (75) yields $\varphi \to \mathcal{A}\varphi$ as a theorem. But we certainly don't want the latter to be a modally-strict theorem; we know that its necessitation, $\Box(\varphi \to \mathcal{A}\varphi)$, fails to be valid – see the discussion following (130.1)★ and, especially, footnote 151. By formulating the consequent of RA with $\mathcal{A}\Gamma$, we forestall such a derivation. All that follows from $\varphi \vdash \varphi$ via RA, as officially formulated in (135), is that $\mathcal{A}\varphi \vdash \mathcal{A}\varphi$, which

[153]Here's why. The justification of RA in the Appendix assumes $\Gamma \vdash \varphi$ and then establishes the conclusion, $\mathcal{A}\Gamma \vdash \mathcal{A}\varphi$, by induction on the length of any derivation that is a witness to $\Gamma \vdash \varphi$. The Base Case, i.e., a derivation of length 1, considers the two ways in which such a derivation could consist of just the single formula φ itself, namely, either φ is an axiom or φ is in Γ. When φ is an axiom, then either (a) it is a necessary axiom or (b) it is either an instance of (43)★ or a universal closure of such an instance. If φ is a necessary axiom, then since the actualizations of all necessary axioms are also axioms, $\mathcal{A}\varphi$ is an axiom, and so a theorem. *A fortiori*, $\mathcal{A}\varphi$ is derivable from $\mathcal{A}\Gamma$. If φ is either an instance of (43)★ or a universal closure of such an instance, then we cite either (133.1) or (134.4) to show $\mathcal{A}\varphi$ is a theorem and thus derivable from $\mathcal{A}\Gamma$.

So if we extend our system with new axioms, the reasoning in the Base Case will be preserved as long as we take their actualizations as axioms or can prove them as theorems.

[154]It is provable that if $\Gamma \models \varphi$, then $\Gamma \models \mathcal{A}\varphi$. Intuitively, if $\Gamma \models \varphi$, i.e., if φ is true at the distinguished world in every interpretation in which all the formulas in Γ are true at the distinguished world, then it follows that in every interpretation in which all the formulas in Γ are true at the distinguished world, $\mathcal{A}\varphi$ is true at the distinguished world, i.e., it follows that $\Gamma \models \mathcal{A}\varphi$.

[155] To see this, assume the antecedent, i.e., $\Gamma \vdash \varphi$. Theorem (130.1)★, which is derived from the modally fragile axiom of actuality (43)★, asserts $\varphi \to \mathcal{A}\varphi$, so that we know $\vdash \varphi \to \mathcal{A}\varphi$. So by (63.10), we have $\varphi \vdash \mathcal{A}\varphi$. But from $\Gamma \vdash \varphi$ and $\varphi \vdash \mathcal{A}\varphi$, it follows by (63.8) that $\Gamma \vdash \mathcal{A}\varphi$. Alternatively, one could justify this rule using an argument similar to that in footnote 156.

Note here how the justification of RA (135) in the Appendix doesn't similarly *use* (43)★. In the base case of the justification, we essentially showed that if φ is any axiom, then $\vdash \mathcal{A}\varphi$, even when φ is axiom (43)★, i.e., even in the case where φ is $\mathcal{A}\psi \to \psi$. Thus, we showed $\vdash \mathcal{A}\varphi$ *without* appealing to (43)★. See the justification of (135) in the Appendix.

by the Deduction Theorem, yields only $\vdash \mathcal{A}\varphi \to \mathcal{A}\varphi$. Moreover, this derivation of $\mathcal{A}\varphi \to \mathcal{A}\varphi$ is modally-strict and we may happily apply RN to derive a valid necessary truth.

Call those rules (like the alternative to RA considered above) whose justification depends on the modally fragile axiom (43)$\star$ *non-strict rules*. If we were to use non-strict rules, we would have to tag any theorem proved by means of such a rule a $\star$-theorem and, indeed, tag the rule itself as a $\star$-rule. This explains why we adopted convention (67.2) described in Remark (67), namely, that we avoid metarules whose justification depends on modally fragile axioms. With such a convention in place, we don't have to worry about redefining *modal strictness* to ensure that derivations that depend on a modally fragile axiom *or* on a non-strict rule fail to be modally-strict.

Third, and finally, note that there are other valid but non-strict metarules that we shall eschew because they violate our convention. Consider, for example, the following rule:

If $\Gamma \vdash \mathcal{A}\varphi$, then $\Gamma \vdash \varphi$.

This rule can be justified from the basis we now have.[156] Again, however, the justification depends on the modally fragile axiom (43)$\star$. Since this reasoning shows us how to turn a proof using the metarule into a proof that doesn't use that rule, it becomes apparent that any proof that uses the above rule implicitly involves an appeal to the modally fragile axiom (43)$\star$. Unless we take further precautions, this rule could permit us to derive invalidities.[157] So instead of taking such precautions as tagging the rule with a $\star$ (to mark it as non-strict) and tagging any derivations involving the rule as non-strict, we simply avoid non-strict rules altogether.

(**137**) **Remark:** The Converse of Weak RN and a Distinction Among Modally Fragile Axioms. The $\star$-axiom $\mathcal{A}\varphi \to \varphi$ (43)$\star$ is traditionally a part of the logic of actuality. It is a distinctive, modally fragile axiom in that its actualization, $\mathcal{A}(\mathcal{A}\varphi \to \varphi)$, is a modally strict theorem (133.4). Given this theorem, an interesting Fact becomes easily provable with respect to the system of axioms Λ defined, in (59), as the axioms asserted in Chapter 8. The Fact is: if φ is a theorem, then $\mathcal{A}\varphi$ is a modally strict theorem:[158]

[156] To see this, assume the antecedent, i.e., $\Gamma \vdash \mathcal{A}\varphi$. Note that since the modally fragile axiom of actuality (43)$\star$ asserts that $\mathcal{A}\varphi \to \varphi$, it follows by (63.1) that $\vdash \mathcal{A}\varphi \to \varphi$. So by (63.3), it follows that $\Gamma \vdash \mathcal{A}\varphi \to \varphi$. Then from $\Gamma \vdash \mathcal{A}\varphi$ and $\Gamma \vdash \mathcal{A}\varphi \to \varphi$, it follows by Modus Ponens that $\Gamma \vdash \varphi$. Alternatively, one could justify this rule using an argument similar to that in footnote 155.

[157] Here is a simple example. As an instance of $\varphi \vdash \varphi$ (63.4), we know: $\mathcal{A}\varphi \vdash \mathcal{A}\varphi$. So the proposed rule of actuality elimination would allow one to infer $\mathcal{A}\varphi \vdash \varphi$. But by the Deduction Theorem (75), it would follow that $\mathcal{A}\varphi \to \varphi$ is a modally-strict theorem. We know that the necessitation of this claim is invalid, but without further constraints on RN, it would follow that $\Box(\mathcal{A}\varphi \to \varphi)$ is a theorem.

[158] The following is stronger than the metarule: if $\vdash \varphi$, then $\vdash \Box\mathcal{A}\varphi$. For this latter metarule is

Fact: If $\vdash \varphi$, then $\vdash_\Box \mathcal{A}\varphi$.

Proof: By induction on the length of a proof of φ.
[The proof is given in the Appendix.]

This Fact will hold as long as our system contains only modally fragile axioms whose actualizations are derivable as modally strict theorems. In the remainder of this Remark, we discuss (a) an interesting simple consequence of this Fact, namely, a proof of the converse of weak RN (i.e., if $\vdash \Box\varphi$, then $\vdash_\Box \varphi$), and (b) a reason why this Fact and the weak converse of RN will fail under natural extensions of the sytem.

Daniel Kirchner and Daniel West have independently pointed out, and proved, that if we limit our system to just the axioms stated in Chapter 8, then the converse of weak RN is indeed valid. Kirchner suggested a proof based on the above Fact:

Converse of Weak RN: If $\vdash \Box\varphi$, then $\vdash_\Box \varphi$.

Proof: Assume $\vdash \Box\varphi$. Then by the Fact just mentioned, $\vdash_\Box \mathcal{A}\Box\varphi$. So by the right-to-left direction of axiom (46.2), $\vdash_\Box \Box\varphi$. And by the T schema, $\vdash_\Box \varphi$.

This nice result suggests that as long as we make sure that any new modally fragile axioms we add to our system have actualizations that are assertible as necessary, the distinction between $\vdash_\Box \varphi$ and $\vdash \Box\varphi$ is not too important. For then both weak RN and converse weak RN are valid, and if we conjoin these conditional metarules, $\vdash_\Box \varphi$ and $\vdash \Box\varphi$ are thereby established as equivalent.

That said, however, it is important to recognize that we may indeed wish to extend our system with modally fragile axioms whose actualizations (a) are not assertible as necessary truths and (b) should not be derivable as modally strict theorems. To see some examples, first consider a scenario we've discussed before, in which we extend our system with the axioms Pa ('a exemplifies P') and $\Diamond\neg Pa$, say where 'a' names some concrete individual and P some property that a exemplifies contingently. Alternatively, suppose we extend our system with the axioms *there exists a unique moon of Earth*, i.e., $\exists!xMxe$, and *Earth might not have had a moon*, i.e., $\Diamond\neg\exists xMxe$. Now in these cases, we would stipulate that Pa and $\exists!xMxe$ are modally fragile and mark them as $\star$-axioms. And to preserve our metarules under such extensions, we would also (a) take the $\Box$-free closures of Pa and $\exists!xMxe$ as modally fragile axioms, and (b) take all the closures of $\Diamond\neg Pa$ and $\Diamond\neg\exists xMxe$ as necessary axioms. Given (a), $\mathcal{A}Pa$ and

easily proved without induction. Assume $\vdash \varphi$. Then RA yields that $\vdash \mathcal{A}\varphi$. But by axiom (46.1) and (63.1), we know $\vdash \mathcal{A}\varphi \rightarrow \Box\mathcal{A}\varphi$. Hence $\vdash \Box\mathcal{A}\varphi$, by (63.6).

$\mathcal{A}\exists!xMxe$ would become asserted as $\star$-axioms.[159] And it would immediately follow, by (46.1), that $\Box\mathcal{A}Pa$ and $\Box\mathcal{A}\exists!xMxe$ become $\star$-theorems.

It is important that we not let this last fact mislead us into thinking that it would be acceptable to assert $\mathcal{A}Pa$ and $\mathcal{A}\exists!xMxe$ as necessary axioms. If we were to do so, then their necessitations (i.e., modal closures) would become necessary axioms as well. Given the understanding of these claims in the previous paragraph, it should *not* be axiomatic that *necessarily, it is actually the case that a exemplifies P* or that *necessarily, it is actually the case that there is a unique moon of the Earth*. Though we can accept $\Box\mathcal{A}Pa$ and $\Box\mathcal{A}\exists!xMxe$ as $\star$-theorems *given* that they become derivable from a contingency, we cannot accept $\Box\mathcal{A}Pa$ and $\Box\mathcal{A}\exists!xMxe$ as necessary axioms. These latter are not unadulterated necessary truths and, should not be axiomatic.

To bring the discussion full circle: if one were to extend our system with modally fragile axioms such as Pa and $\exists!xMxe$, the following would hold:

$\vdash Pa$ and $\nvdash_{\Box} \mathcal{A}Pa$

$\vdash \exists!xMxe$ and $\nvdash_{\Box} \mathcal{A}\exists!xMxe$

So it is possible to extend our system in ways that are inconsistent with the **Fact** (noted at the outset) that currently holds. Pa and $\exists!xMxe$ would be modally fragile axioms but, unlike (43)$\star$, their actualizations would not be modally strict theorems. Thus, there are two kinds of modally fragile axioms: contingent axioms, whether they are known *a posteriori* (such as Pa above) or *a priori* (such as $\mathcal{A}\varphi \rightarrow \varphi$), and necessary truths knowable *a posteriori*, such as $\mathcal{A}Pa$ and $\mathcal{A}\exists!xMxe$. It is extremely important to distinguish them. By designating such necessary *a posteriori* claims as modally fragile axioms, we undermine the proof of the converse of weak RN and immediately obtain a system in which the converse of weak RN fails, as described in Remark (71).

(**138**) $\star$**Theorems:** Actuality and Negation. The following are simple consequences of (130.2)$\star$, and so consequences of (43)$\star$:

(.1) $\neg\mathcal{A}\varphi \equiv \neg\varphi$

(.2) $\neg\mathcal{A}\neg\varphi \equiv \varphi$

Given that the proofs of these theorems depend on (43)$\star$, we may not apply RN to either theorem.

(**139**) **Theorems**. Modally Strict Theorems of Actuality.

(.1) $\mathcal{A}\varphi \lor \mathcal{A}\neg\varphi$

[159]Note how this complies with the requirement that one must the add the actualizations of any new axioms to preserve the Rule of Actualization, as discussed at the end of (135) and in footnote 153.

(.2) $\mathcal{A}(\varphi \,\&\, \psi) \equiv (\mathcal{A}\varphi \,\&\, \mathcal{A}\psi)$

(.3) $\mathcal{A}(\varphi \equiv \psi) \equiv (\mathcal{A}(\varphi \to \psi) \,\&\, \mathcal{A}(\psi \to \varphi))$

(.4) $(\mathcal{A}(\varphi \to \psi) \,\&\, \mathcal{A}(\psi \to \varphi)) \equiv (\mathcal{A}\varphi \equiv \mathcal{A}\psi)$

(.5) $\mathcal{A}(\varphi \equiv \psi) \equiv (\mathcal{A}\varphi \equiv \mathcal{A}\psi)$

(.6) $\mathcal{A}\varphi \equiv \Box\mathcal{A}\varphi$

(.7) $\mathcal{A}\Box\varphi \to \Box\mathcal{A}\varphi$

(.8) $\Box\varphi \to \Box\mathcal{A}\varphi$

(.9) $\mathcal{A}(\varphi \vee \psi) \equiv (\mathcal{A}\varphi \vee \mathcal{A}\psi)$

(.10) $\mathcal{A}\exists\alpha\varphi \equiv \exists\alpha\mathcal{A}\varphi$

(.11) $\mathcal{A}\forall\alpha(\varphi \equiv \psi) \equiv \forall\alpha(\mathcal{A}\varphi \equiv \mathcal{A}\psi)$

Note that one can develop far simpler proofs of some of the above theorems than the ones given in the Appendix by using the modally fragile axiom (43)★. But our policy is to develop modally strict proofs when those are available. Note also that (.2) is used in the proof of (.3), and (.3) and (.4) are used in the proof of (.5). The latter is used to prove (159.1), which is a key lemma in the proof of the Rules of Substitution proved in (160.1) and (160.2).

(**140**) ★**Lemmas:** A Consequence of the Modally Fragile Axiom of Actuality. It is a straightforward consequence of theorem (130.2)★ that an entity α is uniquely such that $\mathcal{A}\varphi$ if and only if α is uniquely such that φ:

> $\forall\beta(\mathcal{A}\varphi^{\beta}_{\alpha} \equiv \beta = \alpha) \equiv \forall\beta(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha)$, provided β is substitutable for α in φ and doesn't occur free in φ

Recall the discussion after we introduced axiom (47): the notation $\mathcal{A}\varphi^{z}_{x}$ used in this theorem involves a harmless ambiguity. Though it should strictly be formulated as $(\mathcal{A}\varphi)^{\tau}_{\nu}$, definition (14) implies $(\mathcal{A}\varphi)^{\tau}_{\nu} = \mathcal{A}(\varphi^{\tau}_{\nu})$.

9.8.2 The Theory of Descriptions

(**141**) ★**Theorems:** Fundamental Theorems Governing Descriptions. It follows from the previous lemma that x is the individual that is (in fact) such that φ just in case x is uniquely such that φ:

> $x = \imath x\varphi \;\equiv\; \forall z(\varphi^{z}_{x} \equiv z = x)$, provided z is substitutable for x in φ and doesn't occur free in φ

As a simple instance, we have $x = \imath xRxa \equiv \forall z(Rza \equiv z = x)$. It follows from this instance by GEN that $\forall x(x = \imath xRxa \equiv \forall z(Rza \equiv z = x))$. So if we instantiate the universal claim to b, we have: $b = \imath xRxa \equiv \forall z(Rza \equiv z = b)$. In other words, b is the object that bears R to a if and only if all and only objects that bear R to a are identical to b.

The proof of the above theorem depends on the axiom for descriptions (47) and lemma (140)★, and hence depends on theorem (130.2)★ and ultimately on the modally fragile axiom for actuality (43)★. So we may not apply RN to this theorem to derive its necessitation. But even though it is a ⋆-theorem, it plays a role in the proof of other important and well-known principles involving descriptions; in the present context, these too are ⋆-theorems. Examples are Hintikka's schema (142)★ and Russell's analysis of descriptions (143)★. Though the classical statement of these well-known principles are derived in a way that is not modally-strict, the principles can be slightly modified so as to be derivable by modally-strict proofs. This will become apparent below.

(**142**) ⋆**Theorems:** Hintikka's Schema. We may derive the instances of Hintikka's schema for definite descriptions namely, x is identical to the individual (in fact) such that φ if and only if φ is true and everything such that φ is identical to x, i.e.,

> $x = \imath x\varphi \equiv (\varphi \,\&\, \forall z(\varphi^z_x \rightarrow z = x))$, provided z is substitutable for x in φ and doesn't occur free in φ

If we apply GEN to the free variable x and then instantiate the result to any variable other than z, say y,[160] then we obtain $y = \imath x\varphi \equiv (\varphi^y_x \,\&\, \forall x(\varphi \rightarrow x = y))$. Cf. Hintikka 1959 (83, 7b), i.e., $\ulcorner y = \imath xf \urcorner \leftrightarrow \ulcorner f(y/x) \,\&\, \forall x(f \rightarrow x = y)\urcorner$. Note that whereas Hintikka's original schema involves a primitive identity symbol, our theorem, which we'll henceforth call Hintikka's schema, involves a defined identity symbol.

The proof of Hintikka's schema appeals to the ⋆-theorem (141)★ and so fails to be modally-strict. Note that in Hintikka's schema, φ is within the scope of the rigidifying operator $\imath x$ on the left side of the biconditional but not within the scope of such an operator on the right side. This is an indicator that (the proof of) the theorem won't be modally strict.

(**143**) ⋆**Theorems:** Russell's Analysis of Descriptions. Our derived quantifier rules also help us to more easily prove, as a *theorem*, a version of Russell's famous (1905) analysis of definite descriptions:

> $\psi^{\imath x\varphi}_x \equiv \exists x(\varphi \,\&\, \forall z(\varphi^z_x \rightarrow z = x) \,\&\, \psi)$, provided (a) ψ is either an exemplification formula $\Pi^n\kappa_1 \ldots \kappa_n$ $(n \geq 1)$ or an encoding formula $\kappa_1 \ldots \kappa_n\Pi^n$

[160] Here, z is the only variable not substitutable for the two free occurrences of x in the above, since z would get captured if we replace the second free occurrence of x.

$(n \geq 1)$, (b) x occurs in ψ and only as one or more of the κ_i $(1 \leq i \leq n)$, and (c) z is substitutable for x in φ and doesn't occur free in φ

This asserts: the individual (in fact) such that φ is such that ψ if and only if something x is such that φ, everything such that φ just is x, and x is such that ψ. Note that Russell's analysis is a $\star$-theorem; the proof relies on Hintikka's schema, which in turn depends on (141)$\star$.

(**144**) $\star$**Theorems:** Significant Descriptions and Uniqueness. Recall again our convention, mentioned at the end of (20), that $\imath x\varphi\downarrow$ abbreviates $(\imath x\varphi)\downarrow$. (We henceforth omit mention of this convention in the remainder of this chapter.) The following are now facts about descriptions and uniqueness: (.1) the x such that φ exists if and only if there exists a unique x such that φ; and (.2) something is identical to the x such that φ if and only if there exists a unique x such that φ. Formally:

(.1) $\imath x\varphi\downarrow \equiv \exists!x\varphi$

(.2) $\exists y(y=\imath x\varphi) \equiv \exists!x\varphi$, provided that y doesn't occur free in φ

Clearly, (.2) follows from (.1) by the instance $\imath x\varphi\downarrow \equiv \exists y(y=\imath x\varphi)$ of theorem (121.1). This instance allows us to regard a description $\imath x\varphi$ as *significant* whenever we know that $\exists y(y=\imath x\varphi)$.

Note that with (.1), we have derived definition $*14{\cdot}02$ in *Principia Mathematica* (Whitehead & Russell 1910–1913 [1925–1927]). On p. 174 of the second edition (p. 182 of the first), we find:

$$*14\cdot 02 \quad E!(\imath x)(\varphi x)\,.\ =:\ (\exists b)\ :\ \varphi x\,.\equiv_x\,.\, x=b \quad \text{Df}$$

To show that (.1) is our version of $*14{\cdot}02$, we need to show that the definiendum and the definiens of the two definitions are equivalent. First, recall again that the discussion on pp. 30–31 ([1925–1927], Introduction, Chapter I) makes it clear that $\imath x\varphi\downarrow$ is the counterpart to their formula $E!(\imath x)(\varphi x)$; their formula $E!(\imath x)(\varphi x)$ asserts that the x such that φ exists. Moreover, our definiens, $\exists!x\varphi$, is equivalent to their definiens, $(\exists b)\ :\ \varphi x\,.\equiv_x\,.\, x=b$, by (127.2), given their notational conventions.

However, since the proofs of our theorems appeal to Hintikka's schema, (.1) and (.2) are both $\star$-theorems and so not subject to RN. The necessitations of both (.1) and (.2) fail to be valid, in both directions.

(**145**) $\star$**Theorems:** Facts About the Matrices of (Significant) Descriptions.

(.1) $x=\imath x\varphi \rightarrow \varphi$

(.2) $z=\imath x\varphi \rightarrow \varphi^z_x$, provided z is substitutable for x in φ

(.3) $\imath x\varphi\downarrow \rightarrow \varphi^{\imath x\varphi}_x$, provided $\imath x\varphi$ is substitutable for x in φ

(.4) $\exists y(y = \imath x\varphi) \rightarrow \varphi_x^{\imath x\varphi}$, provided y doesn't occur free in $\imath x\varphi$ and $\imath x\varphi$ is substitutable for x in φ

These theorems license substitutions into the matrix of a description under certain conditions. Note that in (.2), z may occur free in $\imath x\varphi$. (.3) and (.4) both intuitively tell us we can substitute a significant description into its own matrix. Note that since the proofs appeal to Hintikka's schema, they are all $\star$-theorems.

(**146**) **Lemmas:** Consequence of the Necessary Equivalence of $\mathcal{A}\varphi$ and $\mathcal{A}\mathcal{A}\varphi$. One of the necessary axioms for actuality is (44.4), namely, $\mathcal{A}\varphi \equiv \mathcal{A}\mathcal{A}\varphi$. It is a straightforward, modally-strict consequence of this axiom that an entity α is uniquely such that $\mathcal{A}\varphi$ if and only if α is uniquely such that $\mathcal{A}\mathcal{A}\varphi$:

$\forall\beta(\mathcal{A}\varphi_\alpha^\beta \equiv \beta = \alpha) \equiv \forall\beta(\mathcal{A}\mathcal{A}\varphi_\alpha^\beta \equiv \beta = \alpha)$, provided β is substitutable for α in φ and doesn't occur free in φ

No appeal to the modally fragile axiom for actuality (43)$\star$ is needed to prove this lemma.

(**147**) **Theorems:** Additional Theorems for Descriptions and Actuality. It is provable that (.1) x is identical to the individual such that φ if and only if x is identical to the individual actually such that φ, and (.2) if the x such that φ exists, then it is identical to the x actually such that φ:

(.1) $x = \imath x\varphi \;\equiv\; x = \imath x\mathcal{A}\varphi$

(.2) $\imath x\varphi\downarrow \rightarrow \imath x\varphi = \imath x\mathcal{A}\varphi$

These are modally-strict theorems.

Exercise: Develop modally strict proofs of:

- $\psi_x^{\imath x\varphi} \rightarrow \imath x\mathcal{A}\varphi\downarrow$, provided (a) ψ is either an exemplification formula $\Pi^n\kappa_1 \ldots \kappa_n$ $(n \geq 1)$ or an encoding formula $\kappa_1 \ldots \kappa_n\Pi^n$ $(n \geq 1)$, and (b) x occurs in ψ and only as one or more of the κ_i $(1 \leq i \leq n)$.

- $\psi_x^{\imath x\varphi} \rightarrow \imath x\varphi = \imath x\mathcal{A}\varphi$, provided (a) ψ is either an exemplification formula $\Pi^n\kappa_1 \ldots \kappa_n$ $(n \geq 1)$ or an encoding formula $\kappa_1 \ldots \kappa_n\Pi^n$ $(n \geq 1)$ and (b) x occurs in ψ and only as one or more of the κ_i $(1 \leq i \leq n)$.

(**148**) **Theorems:** Modally Strict Version of Hintikka's Schema. By a judicious placement of the actuality operator, we obtain the following modally strict versions of Hintikka's schema (142)$\star$, namely, x is the individual (in fact) such that φ if and only if it is actually the case that φ and everything actually such that φ is identical to x, i.e.,

$x = \imath x\varphi \equiv \mathcal{A}\varphi \,\&\, \forall z(\mathcal{A}\varphi^z_x \rightarrow z = x)$, provided z is substitutable for x in φ and doesn't occur free in φ.

The proof of this theorem appeals to the necessary axiom (47) instead of to theorem (141)★. Since it depends on no ★-theorems, it is modally-strict.

(**149**) **Theorems:** Descriptions, Actuality, and Necessity. The modally strict version of Hintikka's schema allows us to formulate and prove a nice theorem in connection with the descriptions $\imath x\varphi$ and $\imath x\psi$ whose matrices are (actually or necessarily) universally equivalent. (.1) If it is actually the case that everything is such that φ iff ψ, then for any x, x is identical to the φ iff x is identical to the ψ:

(.1) $\mathcal{A}\forall x(\varphi \equiv \psi) \rightarrow \forall x(x = \imath x\varphi \equiv x = \imath x\psi)$

Note that we can't prove $\mathcal{A}\forall x(\varphi \equiv \psi) \rightarrow \imath x\varphi = \imath x\psi$; the consequent implies that the descriptions are significant, which is something that is not guaranteed by the antecedent. To see this, consider a situation in which both $\neg\exists x\mathcal{A}\varphi$ and $\neg\exists x\mathcal{A}\psi$. Then by (103.9), it would follow that $\forall x(\mathcal{A}\varphi \equiv \mathcal{A}\psi)$ and, by the right-to-left direction of (139.11), $\mathcal{A}\forall x(\varphi \equiv \psi)$, which is the antecedent of (.1). But in this situation, the descriptions $\imath x\varphi$ and $\imath x\psi$ both fail to be significant, and $\imath x\varphi = \imath x\psi$ would be false.

It follows from (.1), by modally strict means, that (.2) if the x such that φ exists and actually everything is such that φ iff ψ, then the x such that φ is identical to the x such that ψ:

(.2) $\imath x\varphi\downarrow \,\&\, \mathcal{A}\forall x(\varphi \equiv \psi) \rightarrow \imath x\varphi = \imath x\psi$

Clearly, then, (.3) if the x such that φ exists and necessarily everything is such that φ iff ψ, then the x such that φ is identical to the x such that ψ:

(.3) $\imath x\varphi\downarrow \,\&\, \Box\forall x(\varphi \equiv \psi) \rightarrow \imath x\varphi = \imath x\psi$

Two other theorems are worthy of mention. (.4) if the x such that φ exists, it does so necessarily, and (.5) if the x such that φ exists, then something is necessarily identical to it:

(.4) $\imath x\varphi\downarrow \rightarrow \Box\imath x\varphi\downarrow$

(.5) $\imath x\varphi\downarrow \rightarrow \exists y\Box(y = \imath x\varphi)$, provided y doesn't occur free in φ

These last two theorems also have modally-strict proofs. The fact that descriptions are rigid is key to understanding both theorems. If we think semantically in terms of a primitive notion of possible world, then we can say that when $\imath x\varphi$ is significant and occurs within the scope of the modal operator $\Box$, it still denotes the object that, *at the actual world*, is uniquely such that φ. Note that

(.4) is simply an instance of (106) and that (.5) follows from the instance of theorem (124.2) that asserts: every individual is necessarily self-identical, i.e., $\forall z\Box(z=z)$.

(**150**) ⋆**Theorems:** Descriptions With Matrices That Obey a Universal Equivalence. We obtain non-modally strict versions of some recent theorems if we start from universalized material equivalences. (.1) If everything is such that φ iff ψ, then for all x, x is identical to the φ if and only if x is identical to the ψ; and (.2) if the x such that φ exists and everything is such that φ iff ψ, then the x such that φ is identical to the x such that ψ:

(.1) $\forall x(\varphi \equiv \psi) \rightarrow \forall x(x=\imath x\varphi \equiv x=\imath x\psi)$

(.2) $\imath x\varphi\downarrow \;\&\; \forall x(\varphi \equiv \psi) \rightarrow \imath x\varphi=\imath x\psi$

(**151**) **Theorems:** Modally Strict Version of Russell's Analysis of Descriptions. By a another judicious placement of the actuality operator, we can prove a modally strict version of Russell's analysis of descriptions. For any exemplification or encoding formula ψ, the individual (in fact) such that φ is such that ψ if and only if something x is such that actually φ, everything such that actually φ just is x, and x is such that ψ, i.e.,

> $\psi^{\imath x\varphi}_x \equiv \exists x(\mathcal{A}\varphi \;\&\; \forall z(\mathcal{A}\varphi^z_x \rightarrow z=x) \;\&\; \psi)$, provided (a) ψ is either an exemplification formula $\Pi^n\kappa_1\ldots\kappa_n$ $(n\geq 1)$ or an encoding formula $\kappa_1\ldots\kappa_n\Pi^n$ $(n\geq 1)$, (b) x occurs in ψ and only as one or more of the κ_i $(1\leq i\leq n)$, and (c) z is substitutable for x in φ and doesn't occur free in φ

No appeal to (43)⋆ or (143)⋆ is necessary.

(**152**) **Theorems:** Theorems for Significant Descriptions and Actuality. We now have:

(.1) $\imath x\varphi\downarrow \equiv \exists!x\mathcal{A}\varphi$

(.2) $x=\imath x\varphi \rightarrow \mathcal{A}\varphi$

(.3) $z=\imath x\varphi \rightarrow \mathcal{A}\varphi^z_x$, provided z is substitutable for x in φ

(.4) $\imath x\varphi\downarrow \rightarrow \mathcal{A}\varphi^{\imath x\varphi}_x$, provided $\imath x\varphi$ is substitutable for x in φ

(.5) $\imath x\varphi=\imath x\psi \rightarrow \mathcal{A}\forall x(\varphi \equiv \psi)$

These are modally-strict theorems. Compare (.1) with (144.1)⋆, and (.2) – (.4) with (145.1)⋆ – (145.3)⋆, respectively.

(**153**) **Theorems:** Modally Strict Conditions For Applying the Matrix of a Description. Here are some interesting facts about descriptions. (.1) If there exists a unique individual that is necessarily such that φ, then anything identical to the x such that φ is such that φ:

(.1) $\exists!x\Box\varphi \rightarrow \forall y(y = \imath x\varphi \rightarrow \varphi^y_x)$, provided y doesn't occur free in φ and is substitutable for x in φ

This is modally strict; by comparing this theorem with (145.2)★, we see that $\exists!x\Box\varphi$ provides modally strict conditions under which φ applies to anything identical to the x such that φ.

Furthermore (.1) helps us to prove another useful theorem, namely, that (.2) if everything such that φ is necessarily such that φ, then if there is a unique thing such that φ, then anything identical to the x such that φ is such that φ:

(.2) $\forall x(\varphi \rightarrow \Box\varphi) \rightarrow (\exists!x\varphi \rightarrow \forall y(y = \imath x\varphi \rightarrow \varphi^y_x))$, provided y doesn't occur free in φ and is substitutable for x in φ

Later in this chapter, we'll see how this theorem helps us to prove facts about a distinguished group of *canonical* abstract objects.

(**154**) **Theorems:** Alphabetically Variant Descriptions. It is relatively straightforward to show that alphabetic variants of significant descriptions can be put into an equation:

$\imath\nu\varphi\downarrow \rightarrow \imath\nu\varphi = (\imath\nu\varphi)'$, where $(\imath\nu\varphi)'$ is any alphabetic variant of $\imath\nu\varphi$

Given the Rule of Alphabetic Variants (114), this theorem is a consequence of the facts (a) that significant descriptions can be instantiated into the universal claim $\forall x(x = x)$ and (b) that the two formulas $\imath\nu\varphi = \imath\nu\varphi$ and $\imath\nu\varphi = (\imath\nu\varphi)'$ are themselves alphabetic variants if $\imath\nu\varphi$ and $(\imath\nu\varphi)'$ are.

(**155**) **Remark:** Digression on Significant Terms. In this Remark, we first produce some examples of complex terms (both descriptions and λ-expressions) that are provably empty and then produce some examples of complex terms that are provably significant despite having provably empty subterms. Specifically, we develop examples of:

(A) descriptions and λ-expressions that are provably empty, i.e., complex terms of the form $\imath z\varphi$ and $[\lambda x\, \psi]$ such that $\vdash \neg\imath z\varphi\downarrow$ and $\vdash \neg[\lambda x\, \psi]\downarrow$,

(B) descriptions that are provably significant despite having a provably empty description as a subterm, i.e., $\imath x\psi$ for which $\vdash \imath x\psi\downarrow$ but where $\imath x\psi$ contains a subterm $\imath z\varphi$ for which $\vdash \neg\imath z\varphi\downarrow$,

(C) descriptions that are provably significant despite having a provably empty λ-expression as a subterm, i.e., $\imath x\psi$ for which $\vdash \imath x\psi\downarrow$ but where $\imath x\psi$ contains a subterm $[\lambda z\, \chi]$ for which $\vdash \neg[\lambda z\, \chi]\downarrow$,

(D) λ-expressions that are provably significant despite having a provably empty description as a subterm; i.e., $[\lambda x\, \psi]$ for which $\vdash [\lambda x\, \psi]\downarrow$ but where $[\lambda x\, \psi]$ contains a subterm $\imath z\varphi$ for which $\vdash \neg\imath z\varphi\downarrow$, and

(E) λ-expressions that are provably significant despite having a provably empty λ-expression as a subterm, i.e., $[\lambda y\, \psi]$ for which $\vdash [\lambda y\, \psi]\downarrow$ but where $[\lambda y\, \psi]$ contains a subterm $[\lambda x\, \chi]$ for which $\vdash \neg[\lambda x\, \chi]\downarrow$.

(A) We begin with examples of a description and a λ-expression that provably fail to be significant. For an empty description, consider $\imath z(Hz \& \neg Hz)$, in which H is a free variable. No matter what property is assigned to the variable H, no object z both exemplifies H and fails to exemplify H. Hence, the object z that both exemplifies H and fails to exemplify H fails to exist; it is an easy theorem that $\neg\imath z(Hz \& \neg Hz)\downarrow$ (exercise). For an empty λ-expression, first recall that in the discussion of the Clark/Boolos paradox in Section 2.1, we sketched how a contradiction follows from the claim $\exists F \forall x(Fx \equiv \exists G(xG \,\&\, \neg Gx))$. This reasoning can be reproduced in our system. Moreover, with the help of the strengthened version of β-Conversion proved in (181) below, we'll develop a formal derivation of $\neg[\lambda x\, \exists G(xG \,\&\, \neg Gx)]\downarrow$ in (192.1).

(B) Next, we give an example of a provably significant description that contains a provably empty one. Since (we've just seen that) $\neg\imath z(Hz \,\&\, \neg Hz)\downarrow$ is a theorem, the contrapositive of axiom (39.5.a) implies that the exemplification formula $G\imath z(Hz \,\&\, \neg Hz)$ is false, i.e., that $\neg G\imath z(Hz \,\&\, \neg Hz)$ is also a theorem, for any property G. Now we shall later prove as a theorem that $y = \imath x(x = y)$ (177.2); intuitively, every individual y is identical to the individual identical to y. Moreover, we leave it as an exercise to show that whenever p is a true proposition, $\imath x(x = y \,\&\, p)$ denotes whatever $\imath x(x = y)$ denotes. Then since $\neg G\imath z(Hz \,\&\, \neg Hz)$ is known to be true, it follows both that:

$$\imath x(x = y) = \imath x(x = y \,\&\, \neg G\imath z(Hz \,\&\, \neg Hz))$$

$$y = \imath x(x = y \,\&\, \neg G\imath z(Hz \,\&\, \neg Hz))$$

for any object y. Since $\imath x(x = y \,\&\, \neg G\imath z(Hz \,\&\, \neg Hz))$ appears in true identity claims, it is significant, i.e.,

(ξ) $\imath x(x = y \,\&\, \neg G\imath z(Hz \,\&\, \neg Hz))\downarrow$,

despite the fact that it contains the provably empty description $\imath z(Hz \,\&\, \neg Hz)$.

(C) Next, we give an example of a provably significant description that contains a provably empty λ-expression. As we've seen, $[\lambda x\, \exists G(xG \,\&\, \neg Gx)]$ is a provably empty λ-expression. So is its alphabetic variant $[\lambda z\, \exists G(zG \,\&\, \neg Gz)]$, by analogous reasoning. Abbreviate this latter λ-expression as $[\lambda z\, \chi]$, so that $\neg[\lambda z\, \chi]\downarrow$ is provable. Then where a is some arbitrarily chosen individual, the formula $\neg[\lambda z\, \chi]a$ is provable, by (39.5.a). Hence, we may formulate the description $\imath x(x = y \,\&\, \neg[\lambda z\, \chi]a)$. Now we've seen that when p is true, $\imath x(x = y \,\&\, p)$ is significant. So $\imath x(x = y \,\&\, \neg[\lambda z\, \chi]a)$ is similarly significant, despite having a provably empty λ-expression as a subterm.

(D) Consider an example discussed in (17.3) and (40). Since we know that $\imath z(Pz \& \neg Pz)$ is a provably empty description, all it takes to see that $[\lambda x Rx\imath z(Pz \& \neg Pz)]$ is a provably significant λ-expression that contains a provably empty description is to note that $[\lambda x\, Rx\imath z(Pz \,\&\, \neg Pz)]\!\downarrow$ is an instance of axiom (39.2).

(E) Finally, we've already seen examples of a provably significant λ-expression that contains a provably empty λ-expression, in the discussions at the end of (39) and (48). For those who skipped those discussions, start with the provably empty λ-expression $[\lambda x\, \exists G(xG \,\&\, \neg Gx)]$. Abbreviate this expression as $[\lambda x\, \chi]$. Since $[\lambda x\, \chi]$ is empty, it follows by (39.5.a) that $\neg[\lambda x\, \chi]y$, for any y. So now consider the λ-expression $[\lambda y\, \neg[\lambda x\, \chi]y]$. This is a core λ-expression and so by axiom (39.2), $[\lambda y\, \neg[\lambda x\, \chi]y]\!\downarrow$. But $[\lambda y\, \neg[\lambda x\, \chi]y]$ contains a provably empty λ-expression as a subterm.

Examples (A) – (E) above show why we may *not* add, as an axiom or hypothesis, the assertion that if a term is significant then all of its subterms are significant, i.e., we may *not* assert:

$$\tau\!\downarrow \rightarrow \sigma\!\downarrow, \text{ whenever } \sigma \text{ is a subterm of } \tau$$

The counterexamples to this claim also explain why we do *not* strengthen (39.5.a) and (39.5.b) to assert that if an n-ary exemplification or encoding formula ($n \geq 1$) is true, then *every* term whatsoever in the formula is significant. To find a counterexample, we construct a true exemplification formula of the form $\Pi\kappa$ such that both primary terms Π and κ contain the provably empty subterm $\imath z(Hz \& \neg Hz)$. First consider the fact that the λ-expression $[\lambda x_1 x_2 \forall F(Fx_1 \equiv Fx_2)]$ is significant, by axiom (39.2). Abbreviate this λ-expression as R, so that we know by β-Conversion (48.2) that $Ryy \equiv \forall F(Fy \equiv Fy)$. But since $\forall F(Fy \equiv Fy)$ is easily established as a theorem, it follows that Ryy. Now consider $[\lambda x\, Rxy]$, i.e., $[\lambda x\, [\lambda x_1 x_2\, \forall F(Fx_1 \equiv Fx_2)]xy]$. This is a core λ-expression and so $[\lambda x\, Rxy]\!\downarrow$, by (39.2). Since β-Conversion now implies $[\lambda x\, Rxy]y \equiv Ryy$, it follows that $[\lambda x\, Rxy]y$. Now recall that in (B) above we established, for any y, that:

$$y = \imath x(x = y \,\&\, \neg G\imath z(Hz \,\&\, \neg Hz))$$

Abbreviate the description as κ, so that the above asserts $y = \kappa$. Then by Rule =E, it follows from $[\lambda x\, Rxy]y$ that $[\lambda x\, Rx\kappa]\kappa$. So where Π is $[\lambda x\, Rx\kappa]$, we have established a true exemplification formula of the form $\Pi\kappa$ in which both Π and κ contain the provably empty description $\imath z(Hz \,\&\, \neg Hz)$. Thus, $\Pi\kappa$ is a true exemplification formula having significant *primary* terms that contain empty subterms.

We leave it as an exercise to show that there are true encoding formulas with primary terms that contain empty subterms. Hence axioms (39.5.a) and (39.5.b) respectively guarantee the significance only of the primary terms in a true exemplification or encoding formula.

9.9 The Theory of Necessity

9.9.1 Propositional Modal Logic

(**156**) **Remark:** Tautologies Are Necessary. In Remark (113), we called attention to Metatheorem ⟨9.2⟩, that every tautology is derivable. We now observe that since the proof of this metatheorem shows that no derivation of a tautology requires an appeal to a modally fragile axiom or a $\star$-theorem, it follows by RN, that every tautology is provably necessary. See Metatheorem ⟨9.3⟩ in the Appendix to this chapter.

(**157**) **Metarules:** Rules RM, RM$\Diamond$, RE, and RE$\Diamond$. The classical metarule RM of modal logic asserts that if $\vdash \varphi \to \psi$, then $\vdash \Box\varphi \to \Box\psi$. However, in our system, Rule RM has to be adjusted slightly to accommodate reasoning with modally fragile axioms and $\star$-theorems.

To see that Rule RM is inappropriate in its classical form, consider what would happen if we were to extend our theory by asserting, for some proposition p, both p and $\neg\Box p$ as axioms. Clearly, p is contingent and since we wouldn't also assert the modal closures of p, p would have to be marked as modally fragile. Now consider the formula $\mathcal{A}p \to p$. Clearly, $\vdash \mathcal{A}p \to p$, since $\mathcal{A}p \to p$ is axiomatic (43)$\star$. By the classical rule RM, it would follow that $\vdash \Box\mathcal{A}p \to \Box p$. But since p is, by hypothesis, an axiom, it is a theorem, i.e., $\vdash p$. So by Rule RA (135), we know $\vdash \mathcal{A}p$. But then by axiom (46.1), it follows that $\vdash \Box\mathcal{A}p$. But this and our previous result (i.e., $\vdash \Box\mathcal{A}p \to \Box p$) yield $\vdash \Box p$. This would mean a contradiction is derivable, since $\neg\Box p$ is also, by hypothesis, an axiom, thereby yielding $\vdash \neg\Box p$.

The problem here is that the classical metarule RM allows one to infer $\vdash \Box\varphi \to \Box\psi$ from any proof of $\varphi \to \psi$, whereas in systems (like the present one) containing modally fragile axioms, it should be restricted so that we may infer $\vdash \Box\varphi \to \Box\psi$ only when there is a modally strict proof of $\varphi \to \psi$. In the previous paragraph, the proof of $\mathcal{A}p \to p$ was not modally strict. If we formulate RM so that it applies to conditional theorems proved by modally strict means, we can forestall the derivation of contradictions in cases analogous to the one just presented. But we first formulate the rule for derivations generally:

(.1) **Rule RM**:
If $\Gamma \vdash_\Box \varphi \to \psi$, then $\Box\Gamma \vdash_\Box \Box\varphi \to \Box\psi$.

Rule RM (Weaker Form):
If $\Gamma \vdash_\Box \varphi \to \psi$, then $\Box\Gamma \vdash \Box\varphi \to \Box\psi$.

In other words, if there is a modally-strict derivation of $\varphi \to \psi$ from Γ, then there is a (modally strict) derivation of $\Box\varphi \to \Box\psi$ from the necessitations of the formulas in Γ. As with RN, we almost always cite the weaker form of RM *with the understanding* that any conclusions drawn via the metarule within a larger reasoning context do not affect the modal strictness of that context (or lack thereof).

When $\Gamma = \varnothing$, then RM reduces to the principle:

- If $\vdash_{\Box} \varphi \to \psi$, then $\vdash_{\Box} \Box\varphi \to \Box\psi$
- If $\vdash_{\Box} \varphi \to \psi$, then $\vdash \Box\varphi \to \Box\psi$ (Weaker Form)

i.e., if $\varphi \to \psi$ is a modally-strict theorem, then $\Box\varphi \to \Box\psi$ is a (modally strict) theorem.

RM$\Diamond$ is a corresponding rule:

(.2) **Rule RM$\Diamond$**:
If $\Gamma \vdash_{\Box} \varphi \to \psi$, then $\Box\Gamma \vdash_{\Box} \Diamond\varphi \to \Diamond\psi$.

Rule RM$\Diamond$ (Weaker Form):
If $\Gamma \vdash_{\Box} \varphi \to \psi$, then $\Box\Gamma \vdash \Diamond\varphi \to \Diamond\psi$.

In other words, if there is a modally-strict derivation of $\varphi \to \psi$ from Γ, then there is a (modally strict) derivation of $\Diamond\varphi \to \Diamond\psi$ from the necessitations of the formulas in Γ. When $\Gamma = \varnothing$, then RM$\Diamond$ reduces to the principle:

- If $\vdash_{\Box} \varphi \to \psi$, then $\vdash_{\Box} \Diamond\varphi \to \Diamond\psi$
- If $\vdash_{\Box} \varphi \to \psi$, then $\vdash \Diamond\varphi \to \Diamond\psi$ (Weaker Form)

i.e., if $\varphi \to \psi$ is a modally-strict theorem, then $\Diamond\varphi \to \Diamond\psi$ is a (modally strict) theorem. Again, as with RN and RM, we almost always cite the weaker form of RM$\Diamond$ *with the understanding* that any conclusions drawn via the metarule within a larger reasoning context do not affect the modal strictness of that context (or lack thereof).

Finally, we have Rules RE and RE$\Diamond$:

(.3) **Rule RE**:
If $\Gamma \vdash_{\Box} \varphi \equiv \psi$, then $\Box\Gamma \vdash_{\Box} \Box\varphi \equiv \Box\psi$.

Rule RE (Weaker Form):
If $\Gamma \vdash_{\Box} \varphi \equiv \psi$, then $\Box\Gamma \vdash \Box\varphi \equiv \Box\psi$.

(.4) **Rule RE$\Diamond$**:
If $\Gamma \vdash_{\Box} \varphi \equiv \psi$, then $\Box\Gamma \vdash_{\Box} \Diamond\varphi \equiv \Diamond\psi$.

Rule RE$\Diamond$ (Weaker Form):
If $\Gamma \vdash_{\Box} \varphi \equiv \psi$, then $\Box\Gamma \vdash \Diamond\varphi \equiv \Diamond\psi$.

If we consider these rules when $\Box\Gamma$ is empty, RE and RE$\Diamond$ become, respectively:

- If $\vdash_\Box \varphi \equiv \psi$, then $\vdash_\Box \Box\varphi \equiv \Box\psi$
  If $\vdash_\Box \varphi \equiv \psi$, then $\vdash \Box\varphi \equiv \Box\psi$ (Weaker Form)
- If $\vdash_\Box \varphi \equiv \psi$, then $\vdash_\Box \Diamond\varphi \equiv \Diamond\psi$
  If $\vdash_\Box \varphi \equiv \psi$, then $\vdash \Diamond\varphi \equiv \Diamond\psi$ (Weaker Form)

Finally, it is important to note the following, when reasoning with a definition by equivalence, i.e., a definition having the form $\varphi \equiv_{df} \psi$. Rule $\equiv$Df of Equivalence by Definition (90.1) tells us that, given such a definition, we know not just that $\vdash \varphi \equiv \psi$ but also $\vdash_\Box \varphi \equiv \psi$ (the $\vdash_\Box$ form of the rule was omitted from (90.1) by convention (67)). So from the definition $\varphi \equiv_{df} \psi$, Rules RE and RE$\Diamond$ immediately allow us to conclude (by modally strict means), respectively, that $\Box\varphi \equiv \Box\psi$, and $\Diamond\varphi \equiv \Diamond\psi$.

(**158**) **Theorems:** Basic K Theorems. The presentation and proofs of some of the following basic theorems that depend upon the K schema have been informed by Hughes & Cresswell 1968 and 1996:

(.1) $\Box\varphi \rightarrow \Box(\psi \rightarrow \varphi)$

(.2) $\Box\neg\varphi \rightarrow \Box(\varphi \rightarrow \psi)$

(.3) $\Box(\varphi \,\&\, \psi) \equiv (\Box\varphi \,\&\, \Box\psi)$

(.4) $\Box(\varphi \equiv \psi) \equiv (\Box(\varphi \rightarrow \psi) \,\&\, \Box(\psi \rightarrow \varphi))$

(.5) $(\Box(\varphi \rightarrow \psi) \,\&\, \Box(\psi \rightarrow \varphi)) \rightarrow (\Box\varphi \equiv \Box\psi)$

(.6) $\Box(\varphi \equiv \psi) \rightarrow (\Box\varphi \equiv \Box\psi)$

(.7) $((\Box\varphi \,\&\, \Box\psi) \vee (\Box\neg\varphi \,\&\, \Box\neg\psi)) \rightarrow \Box(\varphi \equiv \psi)$

(.8) $\Box(\varphi \,\&\, \psi) \rightarrow \Box(\varphi \equiv \psi)$

(.9) $\Box(\neg\varphi \,\&\, \neg\psi) \rightarrow \Box(\varphi \equiv \psi)$

(.10) $\Box\varphi \equiv \Box\neg\neg\varphi$

(.11) $\neg\Box\varphi \equiv \Diamond\neg\varphi$

(.12) $\Box\varphi \equiv \neg\Diamond\neg\varphi$ (Df$\Box$)

(.13) $\Box(\varphi \rightarrow \psi) \rightarrow (\Diamond\varphi \rightarrow \Diamond\psi)$ (K$\Diamond$)

(.14) $\Diamond\Box\varphi \equiv \neg\Box\Diamond\neg\varphi$

(.15) $(\Box\varphi \vee \Box\psi) \rightarrow \Box(\varphi \vee \psi)$

(.16) $(\Box\varphi \,\&\, \Diamond\psi) \rightarrow \Diamond(\varphi \,\&\, \psi)$

We pause here to make a few remarks about (.1) – (.6) only.

If we follow the usual practice of reading the formula $\Box(\chi \to \theta)$ as χ *necessarily implies* θ, then (.1) intuitively guarantees that a necessary truth is necessarily implied by everything, and (.2) intuitively guarantees that a necessary falsehood necessarily implies everything. These facts constituted the "paradoxes of strict implication" (Lewis & Langford 1932 [1959], 511). However, given the meaning of the conditional, they are harmless.[161] (.3) establishes that a necessary conjunction is equivalent to a conjunction of necessities, while (.4) and (.5) are lemmas needed for the proof of (.6), which asserts that the necessity operator distributes over a biconditional.

Note that the converse of (.6), namely $(\Box\varphi \equiv \Box\psi) \to \Box(\varphi \equiv \psi)$, is not a theorem: the material equivalence of $\Box\varphi$ and $\Box\psi$ doesn't imply that the biconditional $\varphi \equiv \psi$ is necessary. To see this, consider an interpretation in which there are two worlds, $\boldsymbol{w}_0$ and $\boldsymbol{w}_1$, such that (a) φ is true at $\boldsymbol{w}_0$ and false at $\boldsymbol{w}_1$, and (b) ψ is false at $\boldsymbol{w}_0$ and true at $\boldsymbol{w}_1$. Then clearly, both $\Box\varphi$ and $\Box\psi$ are false at $\boldsymbol{w}_0$ and so $\Box\varphi \equiv \Box\psi$ is true at $\boldsymbol{w}_0$ (since $\Box\varphi$ and $\Box\psi$ have the same truth value at $\boldsymbol{w}_0$). But the claim $\Box(\varphi \equiv \psi)$ is false at $\boldsymbol{w}_0$: the conditional $\varphi \equiv \psi$ fails at both worlds given that φ and ψ have different truth values at each world. (It is important to be familiar with this counterexample since as we shall see, there are special conditions under which the converse of (.6) holds, namely, if both φ and ψ are necessary when true. See (172.4) and (179.5) below.)

(**159**) **Metarules:** Rules of Necessary Equivalence. Theorems (88.4.b), (88.4.c), (88.4.d), (99.3), (111.5), (139.5), and (158.6) each help to establish one of the cases of the following rule:

(.1) If $\vdash \Box(\psi \equiv \chi)$, then:

(.a) $\vdash \neg\psi \equiv \neg\chi$

(.b) $\vdash (\psi \to \theta) \equiv (\chi \to \theta)$

(.c) $\vdash (\theta \to \psi) \equiv (\theta \to \chi)$

(.d) $\vdash \forall\alpha\psi \equiv \forall\alpha\chi$

(.e) $\vdash [\lambda\, \psi] \equiv [\lambda\, \chi]$

(.f) $\vdash \mathcal{A}\psi \equiv \mathcal{A}\chi$

(.g) $\vdash \Box\psi \equiv \Box\chi$

[161] To see why these results are to be expected, note that the left-to-right direction of (.12) asserts that $\Box\varphi \to \neg\Diamond\neg\varphi$. So, if the negation of a necessary truth φ is not possible, then *a fortiori*, the negation of φ conjoined with any ψ is not possible, i.e., $\Box\varphi \to \neg\Diamond(\psi \,\&\, \neg\varphi)$. But this is provably equivalent to (.1) (exercise). Similarly, the left-to-right direction of (162.1), which is proved later in the text, asserts $\Box\neg\varphi \to \neg\Diamond\varphi$. But if a necessary falsehood φ is not possible, then φ conjoined with the negation of any ψ is not possible, i.e., $\Box\neg\varphi \to \neg\Diamond(\varphi \,\&\, \neg\psi)$. But this last fact is provably equivalent to (.2) (exercise).

It is worth noting here that, by our convention in Remark (67), we omit the presentation of the metarule that results by subtituting $\vdash_\Box$ for $\vdash$ everywhere in (.1). But such a version of (.1) is justified; the justification of (.1) in the Appendix easily converts to a proof of the $\vdash_\Box$ version, since no appeal is made to a modally fragile axiom to justify the metarule.

Since (.1) covers all the cases where a formula ψ (or χ) can occur as a subformula of a formula φ, the following rule is derivable:

(.2) If $\vdash \Box(\psi \equiv \chi)$, then if φ' is the result of substituting the formula χ for zero or more occurrences of ψ where the latter is a subformula of φ, then $\vdash \varphi \equiv \varphi'$.

By convention (67), we omit the formulation and justification of the rule obtained by replacing both occurrences of $\vdash$ by $\vdash_\Box$ in (.2).

A weaker, but much more useful, consequence of (.2) is the following rule: (.3) if $\psi \equiv \chi$ is a modally strict theorem, ψ is a subformula of a formula φ, and φ' is the result of substituting (not necessarily uniformly) χ for the subformula ψ in φ, then it is a theorem that $\varphi \equiv \varphi'$:

(.3) If $\vdash_\Box \psi \equiv \chi$, then if φ' is the result of substituting the formula χ for zero or more occurrences of ψ where the latter is a subformula of φ, then $\vdash \varphi \equiv \varphi'$.

Clearly, the rule obtained by replacing the second occurrence of $\vdash$ by $\vdash_\Box$ in (.3) is also justified, but we leave its formulation and justification to the reader.

(.2) and (.3) are central reasoning principles and the key to the proofs of the Rules of Substitution formulated below. Here are some examples of their use:

Examples of (.2).
If $\vdash \Box(A!x \equiv \neg\Diamond E!x)$, then $\vdash \exists x A!x \equiv \exists x \neg\Diamond E!x$
If $\vdash \Box(Rxy \equiv (Rxy \,\&\, (Qa \vee \neg Qa)))$, then
$\vdash (Pa \,\&\, \Box Rxy) \equiv (Pa \,\&\, \Box(Rxy \,\&\, (Qa \vee \neg Qa)))$

Examples of (.3).
If $\vdash_\Box (A!x \equiv \neg\Diamond E!x)$, then $\vdash \exists x A!x \equiv \exists x \neg\Diamond E!x$
If $\vdash_\Box (Rxy \equiv (Rxy \,\&\, (Qa \vee \neg Qa)))$, then
$\vdash (Pa \,\&\, \Box Rxy) \equiv (Pa \,\&\, \Box(Rxy \,\&\, (Qa \vee \neg Qa)))$

In the second example in (.2) and (.3), we've conjoined a tautology $Qa \vee \neg Qa$ with the formula Rxy and the result is not just necessarily equivalent to Rxy, but also materially equivalent to Rxy by a modally strict proof. Hence we may replace Rxy in the formula, $Pa \,\&\, \Box Rxy$ ($= \varphi$) by the (necessarily and strictly) equivalent formula $Rxy \,\&\, (Qa \vee \neg Qa)$, and the result φ' is (provably) equivalent to φ.

Finally, since a definition-by-equivalence of the form $\psi \equiv_{df} \chi$ yields theorems of the form $\Box(\psi \equiv \chi)$ (among others), such definitions immediately allow us to reason by way of the following rule:

(.4) If $\psi \equiv_{df} \chi$ is a definition-by-equivalence, and φ' is the result of substituting ψ for zero or more occurrences of the χ where the latter is a subformula of φ, then $\vdash \varphi \equiv \varphi'$.

The $\vdash_\Box$ version of this rule is justified by the fact that the definition $\psi \equiv_{df} \chi$ yields *modally strict* theorems of the form $\psi \equiv \chi$. (See the footnote to the proof of (.4) in the Appendix.)

(**160**) **Metarules/Derived Rules:** The Rules of Substitution. The principal Rule of Substitution is (.1) if there is a proof of $\Box(\psi \equiv \chi)$, then in any derivation, ψ and χ can be substituted for one another wherever one or the other occurs as a subformula of any line of the derivation:

(.1) **Rule of Substitution**
If $\vdash \Box(\psi \equiv \chi)$, then where Γ is any set of formulas and φ' is the result of substituting the formula χ for zero or more occurrences of ψ where the latter is a subformula of φ, $\Gamma \vdash \varphi$ if and only if $\Gamma \vdash \varphi'$.
[Variant: If $\vdash \Box(\psi \equiv \chi)$, then where ... , $\varphi \dashv\vdash \varphi'$]

We leave it to the reader to formulate and justify (a) the rule obtained by replacing the unadorned occurrences of $\vdash$ in (.1) by $\vdash_\Box$ and (b) the rule obtained from the Variant by replacing $\vdash$ by $\vdash_\Box$ and $\dashv\vdash$ by ${}_\Box\dashv\vdash_\Box$.

However, a weaker, but much more useful consequence of (.1) is the following, namely, (.2) if there is a modally strict proof of $\psi \equiv \chi$, then in any derivation, ψ and χ can be substituted for one another wherever one or the other occurs as a subformula of any line of the derivation:

(.2) **Rule of Substitution**
If $\vdash_\Box (\psi \equiv \chi)$, then where Γ is any set of formulas and φ' is the result of substituting the formula χ for zero or more occurrences of ψ where the latter is a subformula of φ, $\Gamma \vdash \varphi$ if and only if $\Gamma \vdash \varphi'$.
[Variant: If $\vdash_\Box (\psi \equiv \chi)$, then where ... , $\varphi \dashv\vdash \varphi'$]

Cf. Hughes & Cresswell 1996, 242, **Eq**. We almost always use this form of the rule, since it saves a step; we can reason directly from material equivalences proved by modally strict means, instead of first using Rule RN to show that those equivalences are provable necessary equivalences.

We leave it to the reader to formulate and justify (a) the rule obtained by replacing the two unadorned occurrences of $\vdash$ in (.2) by $\vdash_\Box$ and (b) the rule obtained by replacing $\dashv\vdash$ in the Variant by ${}_\Box\dashv\vdash_\Box$. The justification of these $\vdash_\Box$ versions are straightforward. Since the $\vdash_\Box$ versions are justified, we may cite

(.2) or its Variant within any derivation or proof without having any affect on the modal strictness (or lack thereof) of the reasoning context. That is, we shall, by convention, simply cite the Rule of Substitution with the knowledge that it does not undermine the modal strictness of any reasoning context in which it is applied.

Note that the following is *not* justifiable: if $\vdash \Box(\psi \equiv \chi)$, then $\Gamma \vdash_{\Box} \varphi$ if and only if $\Gamma \vdash_{\Box} \varphi'$. The reason is that if $\vdash \Box(\psi \equiv \chi)$ holds because the only proofs of $\Box(\psi \equiv \chi)$ are $\star$-theorems, then the substitution of χ for ψ in φ to obtain φ' would undermine the modal strictness of the derivation of φ' from Γ.

Finally, since a definition-by-equivalence of the form $\psi \equiv_{df} \chi$ yields necessary axioms, and hence modally strict theorems, of the form $\psi \equiv \chi$, such definitions allow us to reason by way of the following rule:

(.3) **Rule of Substitution for Defined Subformulas**
If $\psi \equiv_{df} \chi$ is a definition-by-equivalence, and φ' is the result of substituting ψ for zero or more occurrences of the χ where the latter is a subformula of φ, then $\Gamma \vdash \varphi$ if and only if $\Gamma \vdash \varphi'$.
[Variant: $\varphi \dashv\vdash \varphi'$, provided . . .]

The $\vdash_{\Box}$ version of (.3) is also easily justified.

(**161**) **Remark:** Digression on Legitimate and Illegitimate Uses of the Rules of Substitution. Here are some legitimate examples of the Variant version of (160.2):

Example 1.
If $\vdash_{\Box} A!x \equiv \neg\Diamond E!x$, then $\neg A!x \dashv\vdash \neg\neg\Diamond E!x$.

Example 2.
If $\vdash_{\Box} Rxy \equiv (Rxy \,\&\, (Qa \vee \neg Qa))$, then $p \rightarrow Rxy \dashv\vdash p \rightarrow (Rxy \,\&\, (Qa \vee \neg Qa))$.

Example 3.
If $\vdash_{\Box} A!x \equiv \neg\Diamond E!x$, then $\exists x A!x \dashv\vdash \exists x \neg\Diamond E!x$.

Example 4.
If $\vdash_{\Box} \neg\neg Px \equiv Px$, then $\mathcal{A}\neg\neg Px \dashv\vdash \mathcal{A}Px$.

Example 5.
If $\vdash_{\Box} (\varphi \rightarrow \psi) \equiv (\neg\psi \rightarrow \neg\varphi)$, then $\Box(\varphi \rightarrow \psi) \dashv\vdash \Box(\neg\psi \rightarrow \neg\varphi)$.

Example 6.
If $\vdash_{\Box} \psi \equiv \chi$, then $\Box(\varphi \rightarrow \psi) \dashv\vdash \Box(\varphi \rightarrow \chi)$.

Example 7.
If $\vdash_{\Box} \varphi \equiv \neg\neg\varphi$, then $\Box(\varphi \rightarrow \varphi) \dashv\vdash \Box(\neg\neg\varphi \rightarrow \varphi)$.

Note, however, that the Rules of Substitution do *not* allow us to substitute χ for ψ in any context whatsoever, but rather only when ψ occurs as a subformula of some given formula:

> **Non-Example 1**. Suppose $\vdash_\Box \psi \equiv \chi$. Then (160.2) does not permit one to substitute χ for ψ in the formula $y[\lambda x\, \psi \,\&\, \theta]$ to infer $y[\lambda x\, \chi \,\&\, \theta]$.

The formula ψ is not a subformula of $y[\lambda x\, \psi \,\&\, \theta]$, and so (160.2) is not applicable.

The following case, in which Γ is empty, also fails to yield a legitimate instance of (160.2):

- **Non-Example 2**. It can be shown that $\vdash_\Box Pa \equiv (Pa \,\&\, (q \vee \neg q))$. But (160.2) does not permit one to substitute $Pa \,\&\, (q \vee \neg q)$ for an occurrence of Pa in the theorem $Pa = Pa$ to infer $Pa = (Pa \,\&\, (q \vee \neg q))$.

Though one can easily establish both $\vdash_\Box Pa \equiv (Pa \,\&\, (q \vee \neg q))$ and $\vdash Pa = Pa$, we may not conclude $\vdash Pa = (Pa \,\&\, (q \vee \neg q))$ because Pa is *not* a subformula of $Pa = Pa$! When one expands the identity symbol in $Pa = Pa$ by its definition, the formula Pa will appear nested with the matrix of complex terms and not as a subformula.[162] Hence, we may not use (160.2) to substitute $Pa \,\&\, (q \vee \neg q)$ for Pa in $Pa = Pa$ to obtain $Pa = (Pa \,\&\, (q \vee \neg q))$.

Similarly, the following is *not* an instance of (160.2):

- **Non-Example 3**. It can be shown that $\vdash_\Box Py \equiv [\lambda x\, Px \,\&\, (q \vee \neg q)]y$. But (160.2) does not permit one to substitute $[\lambda x\, Px \,\&\, (q \vee \neg q)]y$ for an occurrence of Py in the theorem $Py = Py$ to infer $Py = [\lambda x\, Px \,\&\, (q \vee \neg q)]y$.

It is straightforward to establish $\vdash_\Box Py \equiv [\lambda x\, Px \,\&\, (q \vee \neg q)]y$.[163] But Py is not a subformula of the theorem $Py = Py$ and so we can't legitimately apply the rule (160.2) to obtain $Py = [\lambda x\, Px \,\&\, (q \vee \neg q)]y$.

---

[162] By definition (23.4), $Pa = Pa$ is equivalent (eliminating duplicate conjuncts) to:

$$Pa{\downarrow} \,\&\, [\lambda x\, Pa] = [\lambda x\, Pa]$$

This, in turn is equivalent, by definition (23.2), (eliminating duplicate conjuncts) to:

$$(\vartheta)\quad Pa{\downarrow} \,\&\, [\lambda x\, Pa]{\downarrow} \,\&\, \Box\forall y(y[\lambda x\, Pa] \equiv y[\lambda x\, Pa])$$

Note that $Pa{\downarrow}$ expands, by definition (20.3), to $[\lambda x\, Pa]{\downarrow}$. And $[\lambda x\, Pa]{\downarrow}$ (which is also the second conjunct of (ϑ)) expands, by definition (20.2) to $\exists z(z[\lambda x Pa])$. So (ϑ) expands (eliminating duplicate conjuncts) to:

$$(\xi)\quad \exists z(z[\lambda x\, Pa]) \,\&\, \Box\forall y(y[\lambda x\, Pa] \equiv y[\lambda x\, Pa])$$

The definition of *subformula* (6) doesn't count Pa as a subformula of (ξ). The subformulas of the first conjunct are $\exists z(z[\lambda x\, Pa])$ and $z[\lambda x\, Pa]$ and nothing else. As for the second conjunct, $\Box\forall y(y[\lambda x\, Pa] \equiv y[\lambda x\, Pa])$ is a subformula of itself; so by (6.2), $\forall y(y[\lambda x\, Pa] \equiv y[\lambda x\, Pa])$ is also a subformula; by (6.2) again, $y[\lambda x\, Pa] \equiv y[\lambda x\, Pa]$ is also a subformula; and by (6.3) and fact (a) in the final paragraph of (18), $y[\lambda x\, Pa]$ is also a subformula. These are the only subformulas of $\Box\forall y(y[\lambda x\, Pa] \equiv y[\lambda x\, Pa])$.

[163] By (39.2), $[\lambda x\, Px \,\&\, (q \vee \neg q)]{\downarrow}$. So it follows from β-Conversion (48.2) that:

$$[\lambda x\, Px \,\&\, (q \vee \neg q)]y \equiv Py \,\&\, (q \vee \neg q)$$

There are, however, special cases where rule (160.2) can play a useful role legitimizing a substitution that isn't directly allowed by the rule. To see an example, we consider a variant of **Non-Example 1**. Let φ be $[\lambda x\, \psi \,\&\, \theta]y$ instead of $y[\lambda x\, \psi \,\&\, \theta]$. For the reasons discussed earlier, when $\vdash_\Box \psi \equiv \chi$, we may not use (160.2) to substitute χ for ψ in the formula $[\lambda x\, \psi \,\&\, \theta]y$. However, in this particular case, when y is substitutable for x in ψ, χ, and θ, then we can substitute χ for ψ in $[\lambda x\, \psi \,\&\, \theta]y$, though *not solely* by (160.2). That is, the following special case holds:

> **Fact**: If $\vdash_\Box \psi \equiv \chi$, then $[\lambda x\, \psi \,\&\, \theta]y \dashv\vdash [\lambda x\, \chi \,\&\, \theta]y$, provided y is substitutable for x in ψ, χ, and θ.
>
> *Proof*. Suppose $\vdash_\Box \psi \equiv \chi$ and that y is substitutable for x in ψ, χ, and θ. Without loss of generality, we show only $[\lambda x\, \psi \,\&\, \theta]y \vdash [\lambda x\, \chi \,\&\, \theta]y$. Our strategy is to use conditional proof and then (63.10). So assume $[\lambda x\, \psi \,\&\, \theta]y$, to prove $[\lambda x\, \chi \,\&\, \theta]y$. Then $[\lambda x\, \psi \,\&\, \theta]\!\downarrow$, by (39.5.a). Note that from $\vdash_\Box \psi \equiv \chi$, it follows by GEN that $\vdash_\Box \forall x(\psi \equiv \chi)$. Since y is, by hypothesis, substitutable for x in ψ and χ, it is substitutable for x in $\psi \equiv \chi$. So it follows by Rule $\forall$E that $\vdash_\Box \psi^y_x \equiv \chi^y_x$. Note also that if y is substitutable for x in ψ and θ, then y is substitutable for x in $\psi \,\&\, \theta$. Then since $[\lambda x\, \psi \,\&\, \theta]\!\downarrow$, β-Conversion tells us that $[\lambda x\, \psi \,\&\, \theta]y$ implies $(\psi \,\&\, \theta)^y_x$, i.e., $\psi^y_x \,\&\, \theta^y_x$. From this and the previously established fact that $\vdash_\Box \psi^y_x \equiv \chi^y_x$, the Rule of Substitution (160.2) allows us to infer $\chi^y_x \,\&\, \theta^y_x$, i.e., $(\chi \,\&\, \theta)^y_x$. Now if we can show $[\lambda x\, \chi \,\&\, \theta]\!\downarrow$, then we can apply β-Conversion to conclude $[\lambda x\, \chi \,\&\, \theta]y$, thereby completing our conditional proof of $[\lambda x\, \psi \,\&\, \theta]y \to [\lambda x\, \chi \,\&\, \theta]y$; our Fact will then follow by (63.10). But we can show $[\lambda x\, \chi \,\&\, \theta]\!\downarrow$ by first noting that the following is an instance of axiom (49):
>
> $$([\lambda x\, \psi \,\&\, \theta]\!\downarrow \,\&\, \Box\forall x(\psi \,\&\, \theta \equiv \chi \,\&\, \theta)) \to [\lambda x\, \chi \,\&\, \theta]\!\downarrow$$
>
> We already know $[\lambda x\, \psi \,\&\, \theta]\!\downarrow$. So if we can show $\Box\forall x(\psi \,\&\, \theta \equiv \chi \,\&\, \theta)$, then we can conclude $[\lambda x\, \chi \,\&\, \theta]\!\downarrow$ and be done. Note that, by hypothesis, $\vdash_\Box \psi \equiv \chi$. Moreover, we know by (88.4.e) that $\vdash_\Box (\psi \equiv \chi) \to (\psi \& \theta \equiv \chi \& \theta)$. So by (63.6), it follows that $\vdash_\Box \psi \,\&\, \theta \equiv \chi \,\&\, \theta$. It follows by GEN that $\vdash_\Box \forall x(\psi \,\&\, \theta \equiv \chi \,\&\, \theta)$. So by RN, $\vdash \Box\forall x(\psi \,\&\, \theta \equiv \chi \,\&\, \theta)$, which is all that remained for us to show.

So though the Rule of Substitution doesn't by itself let us infer $[\lambda x\, \chi \,\&\, \theta]y$ from $[\lambda x\, \psi \,\&\, \theta]y$ and $\vdash_\Box \psi \equiv \chi$, we can nevertheless draw the inference under certain circumstances.

Now $Py \,\&\, (q \vee \neg q)$ is provably equivalent to Py; it is an instance of the tautology $(\varphi \,\&\, (\psi \vee \neg\psi)) \equiv \varphi$. Thus, it follows by biconditional syllogism and the commutativity of $\equiv$ that $Py \equiv [\lambda x Px \& (q \vee \neg q)]y$. Since this proof is modally-strict, we've established that:

(ϑ) $\vdash_\Box Py \equiv [\lambda x\, Px \,\&\, (q \vee \neg q)]y$

(**162**) **Theorems:** Additional K Theorems. The Rules of Necessary Equivalence and Substitution also make it easier to prove the following theorems:

(.1) $\Box\neg\varphi \equiv \neg\Diamond\varphi$

(.2) $\Diamond(\varphi \vee \psi) \equiv (\Diamond\varphi \vee \Diamond\psi)$

(.3) $\Diamond(\varphi \,\&\, \psi) \rightarrow (\Diamond\varphi \,\&\, \Diamond\psi)$

(.4) $\Diamond(\varphi \rightarrow \psi) \equiv (\Box\varphi \rightarrow \Diamond\psi)$

(.5) $\Diamond\Diamond\varphi \equiv \neg\Box\Box\neg\varphi$

(.6) $\Box(\varphi \vee \psi) \rightarrow (\Box\varphi \vee \Diamond\psi)$

(.7) $(\Box(\varphi \vee \psi) \,\&\, \Diamond\neg\varphi) \rightarrow \Diamond\psi$

(**163**) **Theorems:** The T$\Diamond$ and 5$\Diamond$ Schemata. The T schema (45.2) and the 5 schema (45.3) are axioms of our modal logic. Their duals are theorems:

(.1) $\varphi \rightarrow \Diamond\varphi$ (T$\Diamond$)

(.2) $\Diamond\Box\varphi \rightarrow \Box\varphi$ (5$\Diamond$)

These help us to derive the following classical theorems of propositional S5 modal logic.

(**164**) **Theorems:** Theorems of Actuality, Negation, and Possibility. The Rules of Substitution make it easier to prove (.1) actually φ if and only if it is not the case that actually not φ; (.2) possibly φ if and only if actually possibly φ; (.3) if it is actually the case that φ, then possibly φ; (.4) it is actually the case that φ if and only if it is possible that it is actually the case that φ; and (.5) possibly actually φ implies actually possibly φ:

(.1) $\mathcal{A}\varphi \equiv \neg\mathcal{A}\neg\varphi$

(.2) $\Diamond\varphi \equiv \mathcal{A}\Diamond\varphi$

(.3) $\mathcal{A}\varphi \rightarrow \Diamond\varphi$

(.4) $\mathcal{A}\varphi \equiv \Diamond\mathcal{A}\varphi$

(.5) $\Diamond\mathcal{A}\varphi \rightarrow \mathcal{A}\Diamond\varphi$

Note that the commuted form of (.2), i.e., $\mathcal{A}\Diamond\varphi \equiv \Diamond\varphi$, is another special case of a formula of the form $\mathcal{A}\psi \equiv \psi$ that isn't modally fragile.

(**165**) **Theorems:** Basic S5 Theorems. The following list of basic S5 theorems, i.e., theorems provable from the K, T, and 5 schemata (45.1) – (45.3), was informed by a study of Chellas 1980 (16–18):

(.1) $\Diamond\varphi \equiv \Box\Diamond\varphi$

(.2) $\Box\varphi \equiv \Diamond\Box\varphi$

(.3) $\varphi \to \Box\Diamond\varphi$ (B)

(.4) $\Diamond\Box\varphi \to \varphi$ (B$\Diamond$)

(.5) $\Box\varphi \to \Box\Box\varphi$ (4)

(.6) $\Box\varphi \equiv \Box\Box\varphi$

(.7) $\Diamond\Diamond\varphi \to \Diamond\varphi$ (4$\Diamond$)

(.8) $\Diamond\Diamond\varphi \equiv \Diamond\varphi$

(.9) $\Box(\varphi \vee \Box\psi) \equiv (\Box\varphi \vee \Box\psi)$

(.10) $\Box(\varphi \vee \Diamond\psi) \equiv (\Box\varphi \vee \Diamond\psi)$

(.11) $\Diamond(\varphi \,\&\, \Diamond\psi) \equiv (\Diamond\varphi \,\&\, \Diamond\psi)$

(.12) $\Diamond(\varphi \,\&\, \Box\psi) \equiv (\Diamond\varphi \,\&\, \Box\psi)$

(.13) $\Box(\varphi \to \Box\psi) \equiv \Box(\Diamond\varphi \to \psi)$

(**166**) **Metarules:** Consequences of the B and B$\Diamond$ Schemata. The following rules are derivable with the help of the B (165.3) and B$\Diamond$ (165.4) schemata:

(.1) If $\Gamma \vdash_\Box \Diamond\varphi \to \psi$, then $\Box\Gamma \vdash_\Box \varphi \to \Box\psi$
If $\Gamma \vdash_\Box \Diamond\varphi \to \psi$, then $\Box\Gamma \vdash \varphi \to \Box\psi$ (Weaker Form)

(.2) If $\Gamma \vdash_\Box \varphi \to \Box\psi$, then $\Box\Gamma \vdash_\Box \Diamond\varphi \to \psi$
If $\Gamma \vdash_\Box \varphi \to \Box\psi$, then $\Box\Gamma \vdash \Diamond\varphi \to \psi$ (Weaker Form)

When Γ is empty and there are no premises or assumptions involved, the above reduce to:

- If $\vdash_\Box \Diamond\varphi \to \psi$, then $\vdash_\Box \varphi \to \Box\psi$
  If $\vdash_\Box \Diamond\varphi \to \psi$, then $\vdash \varphi \to \Box\psi$ (Weaker Form)

- If $\vdash_\Box \varphi \to \Box\psi$, then $\vdash_\Box \Diamond\varphi \to \psi$
  If $\vdash_\Box \varphi \to \Box\psi$, then $\vdash \Diamond\varphi \to \psi$ (Weaker Form)

(Cf. Prior 1956, p. 62, Rule RLM.) As with RN, RM, and RE, we almost always cite the weaker versions of these rules, with the understanding that the modal strictness of the reasoning context remains unaffected.

9.9.2 Quantified Modal Logic

(**167**) **Theorems:** Barcan Formulas.

(.1) $\forall\alpha\Box\varphi \to \Box\forall\alpha\varphi$ (Barcan Formula = BF)

(.2) $\Box\forall\alpha\varphi \to \forall\alpha\Box\varphi$ (Converse Barcan Formula = CBF)

(.3) $\Diamond\exists\alpha\varphi \to \exists\alpha\Diamond\varphi$ (BF$\Diamond$)

(.4) $\exists\alpha\Diamond\varphi \to \Diamond\exists\alpha\varphi$ (CBF$\Diamond$)

By an application of RN to (.1), we obtain Theorem 18 of Barcan 1946. By an application of RN to (.3), we obtain Axiom 11 of Barcan 1946.

(**168**) **Theorems:** Other Theorems of Modal Quantification.

(.1) $\exists\alpha\Box\varphi \to \Box\exists\alpha\varphi$ (Buridan)

(.2) $\Diamond\forall\alpha\varphi \to \forall\alpha\Diamond\varphi$ (Buridan$\Diamond$)

(.3) $\Diamond\exists\alpha(\varphi \,\&\, \psi) \to \Diamond(\exists\alpha\varphi \,\&\, \exists\alpha\psi)$

(.4) $\Diamond\exists\alpha(\varphi \,\&\, \psi) \to \Diamond\exists\alpha\varphi$

(.5) $(\Box\forall\alpha(\varphi \to \psi) \,\&\, \Box\forall\alpha(\psi \to \chi)) \to \Box\forall\alpha(\varphi \to \chi)$

(.6) $(\Box\forall\alpha(\varphi \equiv \psi) \,\&\, \Box\forall\alpha(\psi \equiv \chi)) \to \Box\forall\alpha(\varphi \equiv \chi)$

9.9.3 Conditions for, and Consequences of, Modal Collapse

Though it isn't strictly contradictory, a theory of necessity ($\Box$) is nevertheless trivialized if it yields, for an *arbitrary* formula φ, the theorem $\varphi \equiv \Box\varphi$. For if $\varphi \equiv \Box\varphi$ were a theorem for arbitrary φ, then $\neg\varphi \equiv \Box\neg\varphi$ would be an instance, and so by (88.4.b), $\neg\neg\varphi \equiv \neg\Box\neg\varphi$. Thus, by definition of $\Diamond$ and principles of double negation, it would follow that $\varphi \equiv \Diamond\varphi$. So all modal distinctions would disappear; every formula φ would be equivalent to both $\Box\varphi$ and $\Diamond\varphi$. We say that such systems suffer from *modal collapse*.

Clearly, a modally collapsed system is inconsistent with the claim that there is a true proposition that might not have been true (or a false proposition that might have been true).[164] Indeed, modal collapse would produce a contradiction in the present system, since the claims, there are contingent truths and contingent falsehoods, are theorems — see (217.1) and (217.2) below.

[164] Suppose $\exists p(p \,\&\, \Diamond\neg p)$ and let q_0 be such a proposition, so that we know both q_0 and $\Diamond\neg q_0$. If the system is modally collapsed then q_0 implies $\Box q_0$. But $\Diamond\neg q_0$ implies $\neg\Box q_0$. Contradiction.

Analogously, suppose $\exists p(\neg p \,\&\, \Diamond p)$ and let q_1 be such a a proposition, so that we know both $\neg q_1$ and $\Diamond q_1$. If the system is modally collapsed, then as we saw above, $q_1 \equiv \Diamond q_1$ and so $\neg q_1$ implies $\neg\Diamond q_1$. Contradiction.

But though the theory of necessity developed in Sections 9.9.1 and 9.9.2 does not suffer from modal collapse, there are some distinctive kinds of formulas φ for which it is provable that $\Diamond\varphi \equiv \Box\varphi$. We'll therefore say that those formulas exhibit modal collapse. As we shall see, modal collapse for these claims is something that our system embraces to good effect. In what follows, we investigate, in depth, (a) a variety of claims that are modally collapsed, (b) the conditions that give rise to modal collapse, and (c) some consequences of modally collapsed formulas. We'll focus, in turn, on (the conditions for) modal collapse as it affects: existence claims, identity claims, certain modal and actuality claims, claims involving descriptions, and encoding predications. None of these results imply that ordinary (i.e., possibly concrete) individuals are necessarily concrete.

(**169**) **Theorems:** The Modal Collapse of Existence and Nonexistence Claims. We've already seen that $\tau\downarrow \to \Box\tau\downarrow$ is a modally strict theorem (106). The modal logic we now have makes it easy to similarly derive (.1) if it is possible that τ exists, then τ exists:

(.1) $\Diamond\tau\downarrow \to \tau\downarrow$

From (.1) and (106), it immediately follows that (.2) it is possible that τ exists if and only if it is necessary that τ exists:

(.2) $\Diamond\tau\downarrow \equiv \Box\tau\downarrow$

So the existence claim $\tau\downarrow$ is subject to modal collapse. Morever, it is straightforward to show (.3) if τ doesn't exist, then necessarily τ doesn't exist, and (.4) it is possible that τ doesn't exist if and only if it is necessary that τ doesn't exist:

(.3) $\neg\tau\downarrow \to \Box\neg\tau\downarrow$

(.4) $\Diamond\neg\tau\downarrow \equiv \Box\neg\tau\downarrow$

So non-existence claims are subject to modal collapse as well.

(**170**) **Theorems:** Modal Collapse of Identity and Non-identity Claims. We established in (125.1) that $\alpha = \beta \to \Box\alpha = \beta$ is a modally strict theorem. As a consequence, we have that (.1) if α and β are possibly identical then they are identical; (.2) if α and β are distinct, they are necessarily distinct; (.3) if α and β are possibly distinct, they are distinct; (.4) if α and β are possibly distinct, they are necessarily distinct, and (.5) if α and β are possibly distinct, they are necessarily distinct:

(.1) $\Diamond\alpha{=}\beta \to \alpha = \beta$

(.2) $\alpha \neq \beta \to \Box\alpha \neq \beta$

(.3) $\Diamond\alpha \neq \beta \rightarrow \alpha \neq \beta$

(.4) $\Diamond\alpha = \beta \rightarrow \Box\alpha = \beta$

(.5) $\Diamond\alpha \neq \beta \rightarrow \Box\alpha \neq \beta$

In the usual manner, these theorems assume that α and β are both variables of the same type, i.e., both individual variables or both n-ary relation variables, for some n.

(**171**) **Theorems:** The Conditions for Modal Collapse of Complex Formulas. Let χ be any complex formula directly formed from φ (and ψ) by a formation rule (i.e., φ and ψ are immediate subformulas of χ, given the formation rules of the language). Then if φ (and ψ) are modally collapsed then χ is modally collapsed. One subtlety about this fact concerns the case where χ is a universal generalization.

We establish the claim by cases, following the definition of *subformula* in (6), and showing first that it holds when χ is (.1) $\neg\varphi$, (.2) $\varphi \rightarrow \psi$, (.3) $\Box\varphi$, (.4) $\mathcal{A}\varphi$, and (.5) $[\lambda\, \varphi]$:

(.1) $\Box(\varphi \rightarrow \Box\varphi) \rightarrow \Box(\neg\varphi \rightarrow \Box\neg\varphi)$

(.2) $(\Box(\varphi \rightarrow \Box\varphi) \,\&\, \Box(\psi \rightarrow \Box\psi)) \rightarrow \Box((\varphi \rightarrow \psi) \rightarrow \Box(\varphi \rightarrow \psi))$

(.3) $\Box(\varphi \rightarrow \Box\varphi) \rightarrow \Box(\Box\varphi \rightarrow \Box\Box\varphi)$

(.4) $\Box(\varphi \rightarrow \Box\varphi) \rightarrow \Box(\mathcal{A}\varphi \rightarrow \Box\mathcal{A}\varphi)$

(.5) $\Box(\varphi \rightarrow \Box\varphi) \rightarrow \Box([\lambda\, \varphi] \rightarrow \Box[\lambda\, \varphi])$

In the special case where χ is a universal generalization, we then have:

(.6) $\Box\forall\alpha(\varphi \rightarrow \Box\varphi) \rightarrow \Box(\forall\alpha\varphi \rightarrow \Box\forall\alpha\varphi)$

Observe that if $\Box(\varphi \rightarrow \Box\varphi)$ but $\neg\Box\forall\alpha(\varphi \rightarrow \Box\varphi)$, then the modal collapse of φ is spurious, in some sense, since it relies on a particular instantiation to the free variable α. However, if it is a *theorem* that $\Box(\varphi \rightarrow \Box\varphi)$, then by GEN, $\forall\alpha\Box(\varphi \rightarrow \Box\varphi)$, and so by BF (167.1), $\Box\forall\alpha(\varphi \rightarrow \Box\varphi)$. Then from (.6) it follows that $\Box(\forall\alpha\varphi \rightarrow \Box\forall\alpha\varphi)$.

(**172**) **Theorems:** Conditions for, and Consequences of, Modal Collapse. By (165.13), we already know that $\Box(\varphi \rightarrow \Box\varphi)$ is equivalent to $\Box(\Diamond\varphi \rightarrow \varphi)$. We now establish that (.1) necessarily, φ-implies-necessarily-φ if and only if, possibly-φ implies necessarily-φ:

(.1) $\Box(\varphi \rightarrow \Box\varphi) \equiv (\Diamond\varphi \rightarrow \Box\varphi)$

It follows that (.2) if either (a) necessarily, φ-implies-necessarily-φ or (b) possibly-φ implies necessarily-φ, then (c) possibly-φ if and only if necessarily-φ:

(.2) $(\Box(\varphi \to \Box\varphi) \vee (\Diamond\varphi \to \Box\varphi)) \to (\Diamond\varphi \equiv \Box\varphi)$

So if the either condition $\Box(\varphi \to \Box\varphi)$ or $\Diamond\varphi \to \Box\varphi$ holds, $\Diamond\varphi$ and $\Box\varphi$ become equivalent; the distinction between *possibly* φ and *necessarily* φ collapses and thus φ is not subject to modal distinctions.

It also follows that (.3) if necessarily, φ-implies-necessarily-φ, then φ is not necessary if and only if $\neg\varphi$ is necessary:

(.3) $\Box(\varphi \to \Box\varphi) \to (\neg\Box\varphi \equiv \Box\neg\varphi)$

(.3) helps us to establish an important fact about the conditions producing modal collapse, namely, (.4) if necessarily, φ implies $\Box\varphi$, and necessarily, ψ implies $\Box\psi$, then the material equivalence of $\Box\varphi$ and $\Box\psi$ implies that $\varphi \equiv \psi$ is necessary:

(.4) $(\Box(\varphi \to \Box\varphi) \,\&\, \Box(\psi \to \Box\psi)) \to ((\Box\varphi \equiv \Box\psi) \to \Box(\varphi \equiv \psi))$

In other words, (.4) tells us that the converse of (158.6) holds when both φ and ψ are subject to modal collapse. (.4) plays an important role in the proofs of (179.5) and (261.1), which are key theorems.

It follows from the above that (.5) if φ and ψ are modally collapsed, then their material equivalence is modally collapsed:

(.5) $(\Box(\varphi \to \Box\varphi) \,\&\, \Box(\psi \to \Box\psi)) \to \Box((\varphi \equiv \psi) \to \Box(\varphi \equiv \psi))$

Finally, we introduce theorems that prove useful later, namely, (.6) if φ is modally collapsed, then if φ implies that ψ is necessary, then the conditional $\varphi \to \psi$ is necessary; and (.7) if φ is modally collapsed, then if φ implies that ψ is actually true, then the conditional $\varphi \to \psi$ is actually true:

(.6) $\Box(\varphi \to \Box\varphi) \to ((\varphi \to \Box\psi) \to \Box(\varphi \to \psi))$

(.7) $\Box(\varphi \to \Box\varphi) \to ((\varphi \to \mathcal{A}\psi) \to \mathcal{A}(\varphi \to \psi))$

(**173**) **Meta-metarule:** Rule of Modal Strictness. Daniel West has contributed a handy meta-metarule that tells us when a theorem proved by non-modally strict means can in fact be proved by modally strict means. We call this a *meta-metarule* because the rule doesn't govern any formula-forming or term-forming operators, but strictly relates the theorems of two deductive systems $\vdash$ and $\vdash_\Box$, as these notions are defined in (59) and (60). The meta-metarule asserts that if there is proof of a modally collapsed formula φ, then there is a modally strict proof of φ:

Rule of Modal Strictness:
If $\vdash \varphi$ and $\vdash \Box(\varphi \to \Box\varphi)$, then $\vdash_\Box \varphi$

Proof: Assume $\vdash \varphi$ and $\vdash \Box(\varphi \to \Box\varphi)$. Since the T schema (45.2) is an axiom, we also know $\vdash \Box(\varphi \to \Box\varphi) \to (\varphi \to \Box\varphi)$. From this and our second assumption, it follows by (63.5) that $\vdash \varphi \to \Box\varphi$. From this and our first assumption, it similarly follows that $\vdash \Box\varphi$. So by the Converse of Weak RN, which was proved in Remark (137), $\vdash_{\Box} \varphi$.

Note that since we have appealed to the Converse of Weak RN, which holds only in the system as formulated in (59) and (60), the Rule of Modal Strictness is not robust under all extensions of PLM. It will fail in any extension of the system in which we've added modally fragile axioms whose actualizations are also modally fragile; see the discussion in (137). But the rule will hold when we extend the system with any necessary axioms.

This rule is especially useful if one has found a non-modally strict proof of φ and then finds a proof that φ is modally collapsed. The rule then guarantees there is a modally strict proof of φ. By means of this rule, we've been able to find, in earlier drafts of this monograph, theorems that were incorrectly marked as non-modally strict theorems. So it is important to keep the rule in mind when proving what seem to be ★-theorems.

One other interesting consequence of this rule concerns a Rule of Substitution (160.2). Recall that theorem (172.5) tells us that if both ψ and χ are modally collapsed, then $\psi \equiv \chi$ is also modally collapsed. So, any time we are dealing with modally collapsed formulas ψ and χ, we simply have to establish their material equivalence (not necessarily by modally strict means) to use them with the Rule of Substitution. For if we have a proof of $\psi \equiv \chi$ for modally collapsed ψ and χ, then by the Rule of Modal Strictness, we have a modally strict proof of $\psi \equiv \chi$. So by the Rule of Substitution (160.2), one can intersubstitute ψ and χ wherever they occur as subformulas.

(**174**) **Theorems:** Modal Collapse, Actuality, Descriptions, and Unique Existence. From theorem (139.6), i.e., $\mathcal{A}\varphi \equiv \Box\mathcal{A}\varphi$, and theorem (164.4), i.e., $\mathcal{A}\varphi \equiv \Diamond\mathcal{A}\varphi$, it follows that (.1) possibly actually φ if and only if necessarily actually φ:

(.1) $\Diamond\mathcal{A}\varphi \equiv \Box\mathcal{A}\varphi$

Thus, $\mathcal{A}\varphi$ is subject to modal collapse and exhibits no modal distinctions.

Moreover, it follows by modally strict means that if φ is subject to modal collapse, then $\mathcal{A}\varphi$ and φ are provably equivalent, i.e., (.2) if necessarily, if φ then necessarily φ, then $\mathcal{A}\varphi$ if and only if φ:

(.2) $\Box(\varphi \to \Box\varphi) \to (\mathcal{A}\varphi \equiv \varphi)$

Thus, (.2) provides conditions under which $\mathcal{A}\varphi$ and φ become equivalent by modally strict reasoning. Indeed it follows, by modally strict means, that a formula φ is subject to modal collapse if and only if $\mathcal{A}\varphi$ necessarily implies φ:

(.3) $\Box(\varphi \to \Box\varphi) \equiv \Box(\mathcal{A}\varphi \to \varphi)$

The significance of (.3) is that it provides a proof-theoretic justification for prohibiting the application of RN to any instance of axiom (43)★. For (43)★ asserts $\mathcal{A}\varphi \to \varphi$, where this holds for any formula φ. If we could apply RN to conclude $\Box(\mathcal{A}\varphi \to \varphi)$, then by the right-to-left direction of (.3), we could conclude $\Box(\varphi \to \Box\varphi)$, for any formula φ. So an application of RN to (43)★ would lead to the general modal collapse of the system.[165]

There are also results concerning the conditions for modal collapse and definite descriptions. One can prove, by modally strict means, that (.4) if necessarily everything such that φ is necessarily such that φ, then if there is a unique thing such that φ, then the x such that φ exists; and (.5) if necessarily everything such that φ is necessarily such that φ, then x is identical to the x such that φ if and only if φ is true and everything such that φ is identical to x:

(.4) $\Box\forall x(\varphi \to \Box\varphi) \to (\exists!x\varphi \to \imath x\varphi\downarrow)$

(.5) $\Box\forall x(\varphi \to \Box\varphi) \to (x = \imath x\varphi \equiv (\varphi \;\&\; \forall z(\varphi^z_x \to z = x)))$, provided z is substitutable for x in φ and doesn't occur free in φ

By appeal to the Barcan Formula, (.5) implies that if the condition for the modal collapse of φ is universally true, i.e., if $\forall x\Box(\varphi \to \Box\varphi)$, then a modally strict version of the Hintikka scheme (142)★ holds.

Finally, we have (.6) if for every α, φ *necessarily holds* of α whenever it holds of α, then if there is exactly one entity such that φ, then necessarily there is exactly one entity such that φ:

(.6) $\Box\forall\alpha(\varphi \to \Box\varphi) \to (\exists!\alpha\varphi \to \Box\exists!\alpha\varphi)$

(**175**) **Theorems:** Identity and Actuality. The necessity of identity, i.e., $\alpha = \beta \to \Box\alpha = \beta$ (125.1), implies that $\Box(\alpha = \beta \to \Box\alpha = \beta)$, by Rule RN. From this fact, we can easily show that (.1) individuals and n-ary relations ($n \geq 0$) are identical if and only if actually identical, and (2) individuals and n-ary relations are non-identical if and only if actually non-identical:

(.1) $\alpha = \beta \equiv \mathcal{A}\alpha = \beta$, where α, β are variables of the same type

(.2) $\alpha \neq \beta \equiv \mathcal{A}\alpha \neq \beta$, where α, β are variables of the same type

(**176**) **Lemmas:** Actuality and Unique Existence. Since identity claims are modally collapsed, the actuality operator commutes with the unique existence quantifier:

(.1) $\mathcal{A}\exists!\alpha\varphi \equiv \exists!\alpha\mathcal{A}\varphi$

[165] I'm indebted to Daniel West for this observation.

Furthermore, it follows from (.1) and (152.1) that:

(.2) $(\imath x\varphi)\downarrow \equiv \mathcal{A}\exists! x\varphi$

(.2) is especially important. If we can establish a claim of the form $\exists! x\varphi$ by way of a modally strict proof, then the $\vdash_{\Box}$ form of the Rule of Actualization (RA) implies that there is a modally-strict proof of $\mathcal{A}\exists! x\varphi$. Then, by (.2), we can derive $\imath x\varphi\downarrow$ as a modally-strict theorem. This reasoning will play a crucial role when we establish, by modally strict means, that canonical descriptions of the form $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$ are significant (252).

(**177**) **Theorem:** A Distinguished Description Involving Identity. The conditions for the modal collapse of $\alpha = \beta$ were cited in the proof of theorem (175.1). It also plays a key role in the proof of the following theorems, namely, (.1) the individual identical to y exists, and (.2) y is identical to the individual identical to y:

(.1) $\imath x(x = y)\downarrow$

(.2) $y = \imath x(x = y)$

These are modally-strict theorems and though they appears to be trivial, keep in mind that existence and identity are defined notions and that definite descriptions are axiomatized in terms of the actuality operator.

It is worth observing that Frege used a somewhat more sophisticated version of (.2) as an *axiom* governing definite descriptions. In his 1893, he introduces (§11) the function $\backslash\xi$ and uses it to assert Basic Law VI (§18), which we may write as: $y = \backslash\acute{\epsilon}(\epsilon = y)$. In this axiom, the function $\backslash\xi$ has been applied to $\acute{\epsilon}(\epsilon = y)$, which is the extension of a concept under which falls one and only one object, namely, y. Given Frege's explanation of the function $\backslash\xi$ in §11, we can read Law VI as: y is identical to the unique member of the extension of the concept *being identical to y*. By contrast, (.2) is a theorem and asserts that y is identical to the individual identical to y; the description operator isn't applied to an expression for an extension of a concept.

(**178**) **Theorem:** N-ary Encoding and Modality. The modal logic of encoding yields both (.1) if $x_1,\ldots,x_n$ encode F^n, then they necessarily encode F^n, and (.2) if $x_1,\ldots,x_n$ fail to encode F^n, then they necessarily fail to encode F^n:

(.1) $x_1 \ldots x_n F^n \to \Box x_1 \ldots x_n F^n$

(.2) $\neg x_1 \ldots x_n F^n \to \Box\neg x_1 \ldots x_n F^n$

It is relatively easy to derive (.1) from axioms (50) and (51), and the proof of (.2) is simplified by appealing to rule (166.2).

(**179**) **Theorem:** Modal Collapse of Encoding Predications. From the rigidity of unary encoding axiom (51) and the rigidity of n-ary encoding theorem (178.1), we may prove some important *modally-strict* theorems pertaining to the modal collapse of encoding predications, for $n \geq 1$: (.1) possibly $x_1, \ldots, x_n$ encode F^n iff necessarily $x_1, \ldots, x_n$ encode F^n; (.2) $x_1, \ldots, x_n$ encode F^n iff necessarily $x_1, \ldots, x_n$ encode F^n; (.3) possibly $x_1, \ldots, x_n$ encode F^n iff $x_1, \ldots, x_n$ encode F^n; (.4) $x_1, \ldots, x_n F^n$ and $y_1 \ldots y_n G^n$ are materially equivalent iff $\Box x_1, \ldots, x_n F^n$ and $\Box y_1 \ldots y_n G^n$ are materially equivalent; (.5) $x_1, \ldots, x_n F^n$ and $y_1 \ldots y_n G^n$ are necessarily equivalent iff $\Box x_1 \ldots x_n F^n$ and $\Box y_1 \ldots y_n G^n$ are materially equivalent; (.6) $x_1, \ldots, x_n F^n$ and $y_1 \ldots y_n G^n$ are materially equivalent iff it is necessary that they are materially equivalent; (.7) $x_1, \ldots, x_n$ fail to encode F^n iff necessarily $x_1, \ldots, x_n$ fail to encode F^n; (.8) possibly $x_1, \ldots, x_n$ fail to encode F^n iff $x_1, \ldots, x_n$ fail to encode F^n; (.9) possibly $x_1, \ldots, x_n$ fail to encode F^n iff necessarily $x_1, \ldots, x_n$ fail to encode F^n; and (.10) actually $x_1, \ldots, x_n$ encode F^n iff $x_1, \ldots, x_n$ encode F^n:

(.1) $\Diamond x_1 \ldots x_n F^n \equiv \Box x_1 \ldots x_n F^n$ $(n \geq 1)$

(.2) $x_1 \ldots x_n F^n \equiv \Box x_1 \ldots x_n F^n$ $(n \geq 1)$

(.3) $\Diamond x_1 \ldots x_n F^n \equiv x_1 \ldots x_n F^n$ $(n \geq 1)$

(.4) $(x_1 \ldots x_n F^n \equiv y_1 \ldots y_n G^n) \equiv (\Box x_1 \ldots x_n F^n \equiv \Box y_1 \ldots y_n G^n)$ $(n \geq 1)$

(.5) $\Box(x_1 \ldots x_n F^n \equiv y_1 \ldots y_n G^n) \equiv (\Box x_1 \ldots x_n F^n \equiv \Box y_1 \ldots y_n G^n)$ $(n \geq 1)$

(.6) $(x_1 \ldots x_n F^n \equiv y_1 \ldots y_n G^n) \equiv \Box(x_1 \ldots x_n F^n \equiv y_1 \ldots y_n G^n)$ $(n \geq 1)$

(.7) $\neg x_1 \ldots x_n F^n \equiv \Box \neg x_1 \ldots x_n F^n$ $(n \geq 1)$

(.8) $\Diamond \neg x_1 \ldots x_n F^n \equiv \neg x_1 \ldots x_n F^n$ $(n \geq 1)$

(.9) $\Diamond \neg x_1 \ldots x_n F^n \equiv \Box \neg x_1 \ldots x_n F^n$ $(n \geq 1)$

(.10) $\mathcal{A} x_1 \ldots x_n F^n \equiv x_1 \ldots x_n F^n$ $(n \geq 1)$

(.1) is especially significant, given the entanglement of logic and metaphysics. Logically, (.1) is a fact about encoding predications, namely, that they are subject to modal collapse; the possibility and necessity of $x_1 \ldots x_n F^n$ are *equivalent*. Metaphysically, (.1), is a fact about abstract objects, namely, which relations abstract objects possibly encode is not relative to any circumstance.

(.4) – (.6) are also interesting. (.4) tells us that if two encoding formulas are equivalent, then their necessitations are equivalent. (.5) is significant *not* because of the left-to-right direction, which is just an instance (158.6), but because of its right-to-left direction. In general, $\Box\varphi \equiv \Box\psi$ doesn't materially imply the claim $\Box(\varphi \equiv \psi)$, as we saw in the brief discussion following (158.6). But when φ and ψ are two encoding claims, such as $x_1 \ldots x_n F^n$ and $y_1 \ldots y_n G^n$, the implication holds because both encoding claims are subject to modal collapse, as we would expect from the discussion of theorem (172.4). (.6) is a simple but interesting consequence of (.4) and (.5).

Finally, (.10) is significant because it is a modally strict theorem and has the form $\mathcal{A}\varphi \equiv \varphi$, but is provable without an appeal to the non-modally strict theorem schema having this form (130.2)★.

Exercise: Use (.6) to prove $\Diamond(x_1 \ldots x_n F^n \equiv y_1 \ldots y_n G^n) \equiv (x_1 \ldots x_n F^n \equiv y_1 \ldots y_n G^n)$, for $n \geq 1$, and that $\Diamond(x_1 \ldots x_n F^n \equiv y_1 \ldots y_n G^n) \equiv \Box(x_1 \ldots x_n F^n \equiv y_1 \ldots y_n G^n)$.

(**180**) **Theorems:** Being Ordinary and Being Abstract are both Modally Collapsed.

(.1) $O!x \to \Box O!x$

(.2) $A!x \to \Box A!x$

(.3) $\Diamond O!x \to O!x$

(.4) $\Diamond A!x \to A!x$

(.5) $\Diamond O!x \equiv \Box O!x$

(.6) $\Diamond A!x \equiv \Box A!x$

(.7) $O!x \equiv \mathcal{A}O!x$

(.8) $A!x \equiv \mathcal{A}A!x$

The last two theorems are especially interesting. They are commuted claims of the form $\mathcal{A}\varphi \equiv \varphi$ that can be proved by modally-strict means, without an appeal to theorem (130.2)★.

9.10 The Theory of Relations

In this subsection, we describe some important theorems that govern properties, relations, and propositions.

9.10.1 Principles Governing Complex Relation Terms

(**181**) **Theorems:** Strengthened β-Conversion. The axiom β-Conversion (48.2) governs both λ-expressions of the specific form $[\lambda x_1 \ldots x_n\, \varphi]$ and exemplification formulas of the specific form $[\lambda x_1 \ldots x_n\, \varphi]x_1 \ldots x_n$, for $n \geq 1$. But given the Rule of Alphabetic Variants (114), we may now derive a version of β-Conversion for every λ-expression, no matter what distinct individual variables are bound by the λ and no matter what individual variables appear as primary terms in the exemplification formula:

Strengthened β-Conversion: $(n \geq 1)$
$[\lambda\mu_1 \ldots \mu_n\, \varphi]\downarrow \to ([\lambda\mu_1 \ldots \mu_n\, \varphi]\nu_1 \ldots \nu_n \equiv \varphi^{\nu_1,\ldots,\nu_n}_{\mu_1,\ldots,\mu_n})$,
provided $\mu_1, \ldots, \mu_n$ are any distinct individual variables and $\nu_1, \ldots, \nu_n$ are any individual variables substitutable, respectively, for $\mu_1, \ldots, \mu_n$ in φ

The proof is involved and requires that one allow for the presence of free variables in λ-expressions. Note that if we let ν_i be μ_i for all i, $1 \leq i \leq n$, then we have the following:

Special Case: $(n \geq 1)$
$[\lambda\mu_1 \ldots \mu_n\, \varphi]\downarrow \to ([\lambda\mu_1 \ldots \mu_n\, \varphi]\mu_1 \ldots \mu_n \equiv \varphi)$,
provided $\mu_1, \ldots, \mu_n$ are any distinct individual variables.

Clearly, this holds because (a) the definition of *substitutable for* guarantees the μ_i $(1 \leq i \leq n)$ is substitutable for itself in φ, and (b) $\varphi^{\mu_1,\ldots,\mu_n}_{\mu_1,\ldots,\mu_n}$ is just φ.

(**182**) **Remark:** Digression on Free Variables in λ-Expressions. The theorems governing n-ary relations $(n \geq 1)$ developed below are sometimes expressed using λ-expressions that may or may not have free variables. It may prove useful to discuss some examples of such terms. As our first example, consider the following, which is a perfectly good instance of Strengthened β-Conversion, in which two free variables, F and z, occur in the λ-expression:

(ξ) $[\lambda x\, \neg Fxz]\downarrow \to ([\lambda x\, \neg Fxz]y \equiv \neg Fyz)$

This tells us, relative to some F and z, that if $[\lambda x\, \neg Fxz]$ is significant, then an object y exemplifies $[\lambda x\, \neg Fxz]$ if and only if $\neg Fyz$. By GEN, (ξ) implies:

$$\forall y([\lambda x\, \neg Fxz]\downarrow \to ([\lambda x\, \neg Fxz]y \equiv \neg Fyz))$$

and since y doesn't occur free in the antecedent of the conditional, we may, by theorem (95.2), move the quantifier $\forall y$ across the antecedent to conclude that:

(ζ) $[\lambda x\, \neg Fxz]\downarrow \to \forall y([\lambda x\, \neg Fxz]y \equiv \neg Fyz)$

This tells us, relative to some F and z, that if $[\lambda x \neg Fxz]$ is significant, then every object y is such that y exemplifies $[\lambda x\, \neg Fxz]$ if and only if $\neg Fyz$. And since (ζ) was derived from no assumptions, it holds for all F and for all z, by GEN:

$$\forall F \forall z([\lambda x\, \neg Fxz]\downarrow \to \forall y([\lambda x\, \neg Fxz]y \equiv \neg Fyz))$$

Here, we may *not* move the quantifiers $\forall F$ and $\forall z$ across the antecedent, since the variables F and z occur free there. But by theorem (99.14), we may *distribute* the quantifiers $\forall F$ and $\forall z$ over the conditional to conclude:

$$\forall F \forall z([\lambda x\, \neg Fxz]\downarrow) \to \forall F \forall z \forall y([\lambda x\, \neg Fxz]y \equiv \neg Fyz))$$

Note that in this particular example, the antecedent is an axiom; it is a universal closure of an instance of (39.2). Hence, in this example, we can derive a completely general form of β-Conversion, namely:

$$\forall F \forall z \forall y([\lambda x\, \neg Fxz]y \equiv \neg Fyz)$$

in which all the free variables in the matrix $[\lambda x\, \neg Fxz]y \equiv \neg Fyz$ are universally quantified.

Our second example is related to the first; the following is a perfectly good instance of Strengthened β-conversion (181) where, for simplicity, we may take R to be a constant:

$$[\lambda x\, \neg Rxy]\!\downarrow\; \rightarrow ([\lambda x\, \neg Rxy]y \equiv \neg Ryy)$$

This tells us, relative to some object y, that if $[\lambda x\, \neg Rxy]$ is significant, then y exemplifies $[\lambda x\, \neg Rxy]$ if and only if $\neg Ryy$. By GEN, this holds for all y, i.e.,

(ϑ) $\forall y([\lambda x\, \neg Rxy]\!\downarrow\; \rightarrow ([\lambda x\, \neg Rxy]y \equiv \neg Ryy))$

For reasons noted previously, we may *not* move the quantifier across the antecedent to conclude:

$$[\lambda x\, \neg Rxy]\!\downarrow\; \rightarrow \forall y([\lambda x\, \neg Rxy]y \equiv \neg Ryy))$$

But we may distribute the quantifier $\forall y$ in (ϑ) over the conditional to infer:

$$\forall y([\lambda x\, \neg Rxy]\!\downarrow) \rightarrow \forall y([\lambda x\, \neg Rxy]y \equiv \neg Ryy)$$

Again, in this case, the antecedent is an axiom. Consequently, we may derive a general version of β-Conversion for this particular λ-expression, namely:

$$\forall y([\lambda x\, \neg Rxy]y \equiv \neg Ryy)$$

One important moral to draw in these two cases is that it is important to keep track of any variables that occur free in λ-expressions. If the variable α occurs free in $[\lambda x \varphi]$ and $[\lambda x \varphi]\!\downarrow$ is assumed rather than provable as a theorem, we may not infer $\forall \alpha([\lambda x\, \varphi]y \equiv \varphi^y_x)$. If $[\lambda x\, \varphi]\!\downarrow$ is provable as a theorem, then by GEN, it follows that $\forall \alpha([\lambda x\, \varphi]\!\downarrow)$ and in this case, we can then infer $\forall \alpha([\lambda x\, \varphi]y \equiv \varphi^y_x)$.

(**183**) **Theorems:** Corollaries to Strengthened β-Conversion. The foregoing Remark should help one to understand the significance of the following corollaries of strengthened β-Conversion. These hold for any variables meeting the stated conditions:

(.1) $\forall \nu_1 \ldots \forall \nu_n([\lambda \mu_1 \ldots \mu_n\, \varphi]\!\downarrow) \rightarrow \forall \nu_1 \ldots \forall \nu_n([\lambda \mu_1 \ldots \mu_n\, \varphi]\nu_1 \ldots \nu_n \equiv \varphi^{\nu_1,\ldots,\nu_n}_{\mu_1,\ldots,\mu_n})$,
provided $\nu_1, \ldots, \nu_n$ are substitutable, respectively, for $\mu_1, \ldots, \mu_n$ in φ

(.2) $[\lambda \mu_1 \ldots \mu_n\, \varphi]\!\downarrow\; \rightarrow \forall \nu_1 \ldots \forall \nu_n([\lambda \mu_1 \ldots \mu_n\, \varphi]\nu_1 \ldots \nu_n \equiv \varphi^{\nu_1,\ldots,\nu_n}_{\mu_1,\ldots,\mu_n})$,
provided none of $\nu_1, \ldots, \nu_n$ occur free in $[\lambda \mu_1 \ldots \mu_n\, \varphi]$ and $\nu_1, \ldots, \nu_n$ are substitutable, respectively, for $\mu_1, \ldots, \mu_n$ in φ

(**184**) **Derived Rules:** Simple Conditions for Classical β-Conversion. In light of the previous theorems, we formulate simple, *derived* rules of inference for β-Conversion that apply to *closed* λ-expressions only. Let $\kappa_1,\ldots,\kappa_n$ be any individual terms. Then we have:

(.1) **Rule** $\overrightarrow{\beta}$**C**

(.a) $[\lambda\mu_1\ldots\mu_n\,\varphi]\kappa_1\ldots\kappa_n \vdash \varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n}$

(.b) $\neg\varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n} \vdash \neg[\lambda\mu_1\ldots\mu_n\,\varphi]\kappa_1\ldots\kappa_n$

provided $[\lambda\mu_1\ldots\mu_n\,\varphi]$ has no free variables and $\kappa_1,\ldots,\kappa_n$ are any individual terms substitutable, respectively, for $\mu_1,\ldots,\mu_n$ in φ

(.2) **Rule** $\overleftarrow{\beta}$**C**

(.a) $[\lambda\mu_1\ldots\mu_n\,\varphi]\downarrow, \kappa_1\downarrow, \ldots, \kappa_n\downarrow, \varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n} \vdash [\lambda\mu_1\ldots\mu_n\,\varphi]\kappa_1\ldots\kappa_n$

(.b) $[\lambda\mu_1\ldots\mu_n\,\varphi]\downarrow, \kappa_1\downarrow, \ldots, \kappa_n\downarrow, \neg[\lambda\mu_1\ldots\mu_n\,\varphi]\kappa_1\ldots\kappa_n \vdash \neg\varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n}$

provided $[\lambda\mu_1\ldots\mu_n\,\varphi]$ has no free variables and $\kappa_1,\ldots,\kappa_n$ are any individual terms substitutable, respectively, for $\mu_1,\ldots,\mu_n$ in φ

One could try to make these rules stronger, so as to allow for the presence of free variables in $[\lambda\mu_1\ldots\mu_n\,\varphi]$ under certain circumstances. But, in what follows, we won't attempt to strengthen the rules in this way; if we should ever need to reason about λ-expressions with free variables, we'll do so straight from the axioms and theorems that govern them.

(**185**) **Metadefinitions:** η-Variants. We now work our way towards theorems (186) and (187), which are further consequences of η-Conversion (48.3). Our goal over the next few items is to prove that one can always 'reduce' any λ-expression that contains an elementary λ-expression as a subterm. For example, not only want to be able to prove that $[\lambda x\, Fx] = F$, but also the identity claim $[\lambda y\, \neg[\lambda x\, Fx]y] = [\lambda y\, \neg Fy]$. The λ-expressions on the left side of these identity claims are η-reducible. By contrast, it will become clear that a λ-expression such as $[\lambda y\,[\lambda z \neg Pz]y \rightarrow Say]$ is η-irreducible – none of the complex λ-expressions occurring in this example are subject to η-Conversion.

Let ρ be any complex n-ary relation term ($n \geq 0$) so that both ρ and Π range over relation terms. Then we say:

(.1) ρ is *elementary* if and only if ρ has the form $[\lambda\nu_1\ldots\nu_n\,\Pi^n\nu_1\ldots\nu_n]$, where Π^n is any n-ary relation term ($n \geq 0$) and $\nu_1,\ldots,\nu_n$ are distinct individual variables none of which occur free in Π^n.

Note that when $n = 0$, $[\lambda\,\Pi^0]$ is elementary. Furthermore, we say:

(.2) ρ is an *η-expansion of* Π^n if and only if ρ is the elementary λ-expression $[\lambda\nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$

(.3) Π^n is the *η-contraction of* ρ if and only if ρ is the elementary λ-expression $[\lambda\nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$

Note that ρ may have many different η-expansions, depending on the choice of $\nu_1, \ldots, \nu_n$, but an elementary λ-expression ρ can have only one η-contraction.

Now where ρ and ρ' are any n-ary relation terms ($n \geq 0$), we say:

(.4) ρ' is an *immediate* (i.e., one-step) *η-variant of ρ with respect to* Π^n just in case either (a) Π^n is a subterm of ρ and ρ' results from ρ by replacing Π^n by an η-expansion $[\lambda\nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$ or (b) the elementary expression $[\lambda\nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$ is a subterm of ρ and ρ' results from ρ by replacing $[\lambda\nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$ by its η-contraction Π^n.

Clearly, if ρ' is an immediate η-variant of ρ with respect to Π^n, then ρ is an immediate η-variant of ρ' with respect to Π^n. Note that definition (.4) even applies in the case where ρ is φ and ρ' is $[\lambda\,\varphi]$, or vice versa. Thus, φ and $[\lambda\,\varphi]$ are immediate η-variants of each other with respect to φ (which is a 0-ary relation term, according to our BNF). And $[\lambda x\, Px \,\&\, [\lambda\, p]]$ and $[\lambda x\, Px \,\&\, p]$ are immediate η-variants of each other with respect to p.

Finally we say, for n-ary relation terms ρ and ρ':

(.5) ρ' is an *η-variant of* ρ whenever there is a finite sequence of n-ary relation terms $\rho_1, \ldots, \rho_m$ ($m \geq 1$) with $\rho = \rho_1$ and $\rho' = \rho_m$ such that, for every i such that $1 \leq i \leq m-1$, there is some Π^n ($n \geq 0$) such that ρ_{i+1} is an immediate η-variant of ρ_i with respect to Π^n (i.e., such that every member of the sequence is an immediate η-variant of the preceding member of the sequence with respect to some Π^n).

(.6) ρ is *η-irreducible* just in case no relation subterm of ρ is subject to η-contraction (i.e., none of the relation subterms of ρ have an η-contraction).

Thus, the syntactic relation ρ *is an η-variant of* ρ' is the transitive closure of the relation ρ *is an immediate η-variant of* ρ'.[166] We now illustrate these definitions with examples:

η-Variants of Elementary λ-Expressions (Expansion/Contraction Pairs):

- $[\lambda xyz\, F^3xyz]$ / F^3

[166] Intuitively, the transitive closure of R is that relation R' that relates any two elements in a chain of R-related elements. That is, for any elements x and y of a domain of R-related elements, $xR'y$ holds whenever there exist $z_0, z_1, \ldots, z_n$ such that (i) $z_0 = x$, (ii) $z_n = y$, and (iii) for all $0 \leq i < n$, $z_i R z_{i+1}$.

- $[\lambda x\, [\lambda y\, \neg Fy]x]$ / $[\lambda y\, \neg Fy]$
- $[\lambda xy\, [\lambda uv\, \Box\forall F(Fu \equiv Fv)]xy]$ / $[\lambda uv\, \Box\forall F(Fu \equiv Fv)]$
- $[\lambda\, p]$ / p
- $[\lambda\, \neg Pa]$ / $\neg Pa$

Immediate η-Variants

- All of the above
- $[\lambda y\, [\lambda z\, Pz]y \rightarrow Say]$ / $[\lambda y\, Py \rightarrow Say]$
- $[\lambda y\, Py \rightarrow [\lambda uv\, Suv]ay]$ / $[\lambda y\, Py \rightarrow Say]$
- $[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay]$ / $[\lambda y\, [\lambda z\, Pz]y \rightarrow Say]$
- $[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay]$ / $[\lambda y\, Py \rightarrow [\lambda uv\, Suv]ay]$
- $[\lambda y\, [\lambda\, p]]$ / $[\lambda y\, p]$
- $[\lambda x_1 \ldots x_n\, [\lambda\, Pa]]$ / $[\lambda x_1 \ldots x_n\, Pa]$
- $[\lambda z\, Pz]y \rightarrow Say$ / $Py \rightarrow Say$
- $[\lambda\, [\lambda z\, Pz]y]$ / $[\lambda\, Py]$

η-Variant Pairs

- All of the above
- $[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay]$ / $[\lambda y\, Py \rightarrow Say]$
- $[\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay$ / $Py \rightarrow Say$
- $[\lambda y\, [\lambda z\, Pz]y \rightarrow Say]$ / $[\lambda y\, Py \rightarrow [\lambda uv\, Suv]ay]$

With a thorough grip on the notion of η-variants, we may prove the following theorems. Note: the above examples are simple examples since none of the λ-expressions have non-denoting subterms.

(**186**) **Lemmas:** Useful Facts About η-Conversion. We've already proved, for 0-ary relation terms, that η-Conversion holds unconditionally; theorem (111.1) asserts $[\lambda\, \varphi] = \varphi$. But for n-ary relation terms where $n \geq 1$, η-Conversion is conditional on the significance of the relation term to be identified with its η-expansion:

(.1) $\Pi^n\!\downarrow\, \rightarrow [\lambda x_1 \ldots x_n\, \Pi^n x_1 \ldots x_n] = \Pi^n$, where $x_1, \ldots, x_n$ are distinct individual variables and Π^n is any n-ary relation term ($n \geq 0$) in which none of $x_1, \ldots, x_n$ occur free

A somewhat more general form of (.1) holds for elementary λ-expressions constructed with any individual variables $\nu_1, \ldots, \nu_n$, not just those contructed with $x_1, \ldots, x_n$:

(.2) $\Pi^n\downarrow \rightarrow [\lambda\nu_1 \ldots \nu_n\, \Pi^n\nu_1 \ldots \nu_n] = \Pi^n$, where $\nu_1,\ldots,\nu_n$ are any distinct individual variables and Π^n is any n-ary relation term ($n \geq 0$) in which none of $\nu_1,\ldots,\nu_n$ occur free

So while $[\lambda x\, \exists G(xG \,\&\, \neg Gx)]\downarrow \rightarrow ([\lambda z\, [\lambda x\, \exists G(xG \,\&\, \neg Gx)]z] = [\lambda x\, \exists G(xG \,\&\, \neg Gx)])$ is an instance of (.2), we can't detach the consequent because it is provable that $\neg[\lambda x\, \exists G(xG \,\&\, \neg Gx)]\downarrow$ (192.1).

(**187**) **Metarule:** Conditions Under Which η-Conversion holds for Significant η-Variants. Where ρ is any complex relation term, the most general form of η–Conversion is the following:

> **Rule of η-Conversion** (ηC)
> If (a) $\vdash \rho\downarrow$, (b) ρ' is an η-variant of ρ witnessed by the sequence $\rho_1,\ldots,\rho_m$, for which $\rho = \rho_1$ and $\rho' = \rho_m$ ($m \geq 1$), and (c) $\vdash \Pi^n\downarrow$ ($n \geq 0$) whenever ρ_{i+1} is an immediate η-variant of ρ_i with respect to Π^n, for each i such that $1 \leq i \leq m-1$, then $\vdash \rho = \rho'$.

This metarule allows us to collapse all the η-variants within a complex λ-expression in a single stroke, provided that each of the elementary λ-expressions involved is being collapsed to a significant relation term. It allows us to give a handy, 2-line proof of the equation between $[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay]$ and its η-irreducible form $[\lambda y\, Py \rightarrow Say]$:

(a)	$[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay]\downarrow$	(39.2)
(b)	$[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay] = [\lambda y\, Py \rightarrow Say]$	ηC, (a)

This is an instance of the rule because $\vdash P\downarrow$ and $\vdash S\downarrow$ and so the collapse of $[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay]$ to $[\lambda y\, Py \rightarrow Say]$ goes by way of the sequence $[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay]$, $[\lambda y\, Py \rightarrow [\lambda uv\, Suv]ay]$, $[\lambda y\, Py \rightarrow Say]$. Each adjacent pair of expressions in the sequence satisfies clause (c) of the rule.

So, under the conditions stated by the metarule, we may *collapse* some η-variants to their η-irreducible forms to simplify formulas.

(**188**) **Theorems:** Relation Terms That Differ by Co-Denoting Descriptions. It should be clear why we state these next theorems in the formal mode. (.1) If a significant λ-expression contains the description $\imath x\varphi$, and $\imath x\varphi = \imath x\psi$, then the result of substituting $\imath x\psi$ for one or more occurrences of $\imath x\varphi$ in the original λ-expression yields a λ-expression which can be put into an equation with the original; (.2) if $\imath x\varphi = \imath x\psi$, then the result of substituting $\imath x\psi$ for one or more occurrences of $\imath x\varphi$ in a 0-ary relation term Π containing $\imath x\varphi$ yields a 0-ary relation term Π' such that $\Pi = \Pi'$:

(.1) $[\lambda z_1 \ldots z_n\, \chi^{\imath x\varphi}_y]\downarrow \,\&\, \imath x\varphi = \imath x\psi \,\rightarrow\, [\lambda z_1 \ldots z_n\, \chi^{\imath x\varphi}_y] = [\lambda z_1 \ldots z_n\, \chi']$, provided $\imath x\varphi$ and $\imath x\psi$ are both substitutable for y in $[\lambda z_1 \ldots z_n\, \chi]$ and χ' is the result of substituting $\imath x\psi$ for one or more occurrences of $\imath x\varphi$ in $\chi^{\imath x\varphi}_y$ ($n \geq 1$)

(.2) $\imath x\varphi = \imath x\psi \rightarrow \Pi_y^{\imath x\varphi} = \Pi'$, provided Π is any 0-ary relation term, $\imath x\varphi$ and $\imath x\psi$ are both substitutable for y in Π, and Π' is the result of substituting $\imath x\psi$ for one or more occurrences of $\imath x\varphi$ in $\Pi_y^{\imath x\varphi}$.

Intuitively, these theorems tell us that when we have two descriptions that denote the same object, then any two relation terms that differ only by substituting one description for the other (not necessarily uniformly) can be put into an equation.

Some examples might prove useful. From the fact that the property *lands on the moon of Earth* exists ($[\lambda z\, Lz\imath xMxe]\downarrow$) and the fact that the moon of Earth is identical to the moon of the planet inhabited by *homo sapiens* ($\imath xMxe = \imath y(My\imath z(Pz \,\&\, Izh))$), (.1) allows us to infer that the property *lands on the moon of Earth* is identical to the property *lands on the moon of the planet inhabited by* homo sapiens, i.e., $[\lambda z\, Lz\imath xMxe] = [\lambda z\, Lz\imath y(My\imath z(Pz \,\&\, Izh))]$). From $\imath xMxe = \imath y(My\imath z(Pz \& Izh))$, it follows by (.2) that the proposition *that Armstrong lands on the moon of Earth* ($[\lambda\, La\imath xMxe]$) is equal to the proposition *that Armstrong lands on the moon of the planet inhabited by* homo sapiens ($[\lambda\, La\imath y(My\imath z(Pz \,\&\, Izh))]$).

It is important to remember here that the formal representations just offered are not intended to be sensitive to the cognitive significance or Fregean sense of the English noun phrases 'moon of Earth' and 'moon of the planet inhabited by *homo sapien*'. *Mutatis mutandis*, they are not intended to be sensitive to the cognitive significance or Fregean sense of the descriptions 'the moon of the Earth' and 'the moon of the planet inhabited by *homo sapiens*', nor to be sensitive to the cognitive significance or Fregean sense of the verb phrases 'landed on the moon of Earth' and 'landed on the moon of the planet inhabited by *homo sapiens*'. So our formal representations are not intended to represent how these expressions function in intensional contexts like "*S* believes that Armstrong landed on the moon of Earth" and "*S* believes that Armstrong landed on the moon of the planet inhabited by *homo sapiens*". Our formal representations of these inferences capture the *de re* reading on which *only the denotation* of a (significant, natural language) description is relevant to both (a) the identity of the property denoted by a noun or verb phrase in which that description occurs, and (b) the identity of the proposition denoted by a sentence in which that description occurs. Typed object theory is needed to analyze the Fregean sense of a noun or verb phrase and explain why distinct such phrases can signify the same entity but have different senses (and hence distinct cognitive values). Consequently, we won't pursue this issue any further here; see Zalta 1988 (Part IV).

9.10.2 Facts About Relations

(**189**) **Theorem:** Equivalence and Identity of Properties. The following theorem establishes that our theory of properties constitutes an *extensional* theory of hyperintensional entities, since it asserts that properties (which are more fine-grained than intensions classically understood, i.e., more fine-grained than functions from possible worlds to sets of individuals) are identical whenever they are materially equivalent in the encoding sense:

$$F^1 = G^1 \equiv \forall x(xF^1 \equiv xG^1)$$

The proof of this theorem is relatively easy, given theorems (179.2) and (179.5). The left-to-right direction is trivial, but for the right-to-left direction, one can straightforwardly show that $\forall x(xF^1 \equiv xG^1)$ implies $\Box\forall x(xF^1 \equiv xG^1)$, which by theorem (116.1), implies $F^1 = G^1$.[167]

The important point here, then, is that when we have to prove that properties F and G are identical, we need not prove that they are necessarily encoded by the same objects, but only that they are encoded by the same objects.

(**190**) **Remark:** A Digression on the Identity of Relations. Our theory of identity for n-ary relations now consists of the following principles (some of which are simplifications of others):

- $F = G \equiv_{df} F{\downarrow} \,\&\, G{\downarrow} \,\&\, \Box\forall x(xF \equiv xG)$ (23.2)
- $F = G \equiv_{df} F{\downarrow} \,\&\, G{\downarrow} \,\&$
  $\forall y_1 \ldots \forall y_{n-1}([\lambda x\, Fxy_1 \ldots y_{n-1}] = [\lambda x\, Gxy_1 \ldots y_{n-1}] \,\&$
  $[\lambda x\, Fy_1xy_2 \ldots y_{n-1}] = [\lambda x\, Gy_1xy_2 \ldots y_{n-1}] \,\&\ldots\&$
  $[\lambda x\, Fy_1 \ldots y_{n-1}x] = [\lambda x\, Gy_1 \ldots y_{n-1}x])$ (23.3)
- $p = q \equiv_{df} p{\downarrow} \,\&\, q{\downarrow} \,\&\, [\lambda x\, p] = [\lambda x\, q]$ (23.4)
- $[\lambda x_1 \ldots x_n\, \varphi]{\downarrow} \to [\lambda x_1 \ldots x_n\, \varphi] = [\lambda x_1 \ldots x_n\, \varphi]'$,
  where $[\lambda x_1 \ldots x_n\, \varphi]'$ is any alphabetic variant of $[\lambda x_1 \ldots x_n\, \varphi]$ (48.1)
- $[\lambda x_1 \ldots x_n\, F^n x_1 \ldots x_n] = F^n$ (48.3)
- $F = G \equiv \Box\forall x(xF \equiv xG)$ (116.1)

[167] Note that we cannot generalize the proof of the right-to-left direction, i.e., we cannot generalize the proof of $\forall x(xF^1 \equiv xG^1) \to F^1 = G^1$ to establish $\forall x_1 \ldots \forall x_n(x_1 \ldots x_n F^n \equiv x_1 \ldots x_n G^n) \to F^n = G^n$ for $n \geq 2$. Though one can show:

$$\forall x_1 \ldots \forall x_n(x_1 \ldots x_n F^n \equiv x_1 \ldots x_n G^n) \to \Box\forall x_1 \ldots \forall x_n(x_1 \ldots x_n F^n \equiv x_1 \ldots x_n G^n)$$

this isn't sufficient for n-ary relation identity when $n \geq 2$. For (23.3) requires us to show that each way of projecting F^n and G^n onto $n-1$ objects yields identical properties.

- $F^n = G^n \equiv \quad (n \geq 2)$
  $\forall y_1 \dots \forall y_{n-1}([\lambda x\, F^n x y_1 \dots y_{n-1}] = [\lambda x\, G^n x y_1 \dots y_{n-1}]\ \&$
  $[\lambda x\, F^n y_1 x y_2 \dots y_{n-1}] = [\lambda x\, G^n y_1 x y_2 \dots y_{n-1}]\ \& \dots \&$
  $[\lambda x\, F^n y_1 \dots y_{n-1} x] = [\lambda x\, G^n y_1 \dots y_{n-1} x])$ (116.2)
- $p = q \equiv [\lambda y\, p] = [\lambda y\, q]$ (116.3)
- $p = q \equiv [\lambda\, p] = [\lambda\, q]$ (Exercise)
- $[\lambda\, \varphi] = \varphi$ (111.1)
- $[\lambda\, \varphi] = [\lambda\, \varphi]'$, where $[\lambda\, \varphi]$ and $[\lambda\, \varphi]'$ are alphabetic variants (111.3)
- $\varphi = \varphi'$, where φ and φ' are alphabetic variants (111.4)
- $\Pi^n\!\downarrow \to [\lambda \nu_1 \dots \nu_n\, \Pi^n \nu_1 \dots \nu_n] = \Pi^n$, where $\nu_1, \dots, \nu_n$ are any distinct individual variables and Π^n is any n-ary relation term ($n \geq 0$) in which none of $\nu_1, \dots, \nu_n$ occur free (186.2)
- Rule ηC (187)
- $F = G \equiv \forall x(xF \equiv xG)$ (189)

Each of these principles provides a perspective on the question: under what conditions are relations identical?

Note that these principles do *not* constitute a procedure for determining whether an arbitrary pair of (nominalized) predicates of natural language denote the same property or relation, nor a procedure for determining whether the members of such pairs denote at all. We've seen, for example, that the theory does *not* predict that the property *being a woodchuck* is the same property as *being a groundhog*. This is a theoretical identity that can't be established *a priori*. Similarly, if we represent *being red and round* as $[\lambda x\, R_1 x \,\&\, R_2 x]$, represent *being round and red* as $[\lambda x\, R_2 x \,\&\, R_1 x]$, then although our theory tells us that both formal expressions are significant, it doesn't force these expressions to denote the same property; it leaves one free to assert or deny this particular identity. And should one believe there is an over-riding reason to do so, one could assert $\forall F \forall G([\lambda x\, Fx \,\&\, Gx] = [\lambda x\, Gx \,\&\, Fx])$ and other, similar principles.

(**191**) **Theorems:** Comprehension Principles for Relations and Properties. The following is a theorem schema derivable directly from the axiom of β-Conversion (48.2) and constitutes a comprehension principle for relations:

(.1) $\exists F^n \Box \forall x_1 \dots \forall x_n (F^n x_1 \dots x_n \equiv \varphi)$, where $n \geq 1$, F^n doesn't occur free in φ, and none of $x_1, \dots, x_n$ occur in encoding position in φ.

When $n = 1$, (.1) becomes a comprehension principle for properties:

(.2) $\exists F\Box\forall x(Fx \equiv \varphi)$, provided F doesn't occur free in φ and x doesn't occur free in encoding position in φ.

We'll derive a comprehension principle for propositions a bit later, in (194).

(**192**) **Theorems:** The Paradoxical λ-Expressions Don't Denote Relations. It is important to see that the λ-expression that leads to the Clark/Boolos paradox is provably empty:

(.1) $\neg[\lambda x\, \exists G(xG \,\&\, \neg Gx)]\downarrow$

We also know that (.2) there is no property that is exemplified by all and only the objects x that encode a property that x fails to exemplify:

(.2) $\neg\exists F\forall x(Fx \equiv \exists G(xG \,\&\, \neg Gx))$

Moreover, our system implies that the λ-expression which leads to the McMichael-Boolos paradox, namely, $[\lambda z\, z\!=\!y]$, doesn't denote a property for every object y:

(.3) $\neg\forall y([\lambda z\, z\!=\!y]\downarrow)$

Note that this does *not* assert that $\forall y\neg([\lambda z\, z\!=\!y]\downarrow)$. We'll prove later that if y is ordinary, *being identical to* y exists, i.e., $O!y \rightarrow [\lambda z\, z\!=\!y]\downarrow$. See theorem (240.2).

Consequently, our system implies (.4) it is not the case that for every object y, there exists a property exemplified by all and only the individuals identical to y:

(.4) $\neg\forall y\exists F\forall x(Fx \equiv x\!=\!y)$

Clearly, it also follows that it is not the case that there is a relation F that objects x and y exemplify whenever they are identical:

(.5) $\neg\exists F\forall x\forall y(Fxy \equiv x\!=\!y)$

We'll see later, however, in (229), that the relation, *being an x and y such that x and y are identical ordinary objects*, exists. Thus, there are identity *relations* that hold among the objects of certain restricted subdomains. We'll see several other such relations later in this text.

(**193**) ⋆**Theorem:** The Kirchner Paradox is Blocked. (.1) If everything is G, then *being an x such that the individual y, which is both identical to x and such that x fails to exemplify a property it encodes, exemplifies G* doesn't exist:

(.1) $\forall xGx \rightarrow \neg[\lambda x\, G\iota y(y\!=\!x \,\&\, \exists H(xH \,\&\, \neg Hx))]\downarrow$

Since there are properties that are universal, we can conclude that (.2) there are properties for which the λ-expressions leading to the Kirchner paradox fail to be significant:

(.2) $\exists G(\neg[\lambda x\, G\imath y(y=x \,\&\, \exists H(xH \,\&\, \neg Hx))]\downarrow)$

Exercise: Develop and justify answers to the questions: Are there modally strict proofs of (.1) and (.2)?

(**194**) **Theorems:** Comprehension Principle for Propositions. From the fact that every formula is a term that is provably significant (104.2), we obtain the following comprehension principle for propositions:

$\exists p\Box(p \equiv \varphi)$,
where φ is any formula with no free occurrences of p.

This comprehension principle and theorem (116.2) jointly offer a precise theory of propositions. On this theory, the claim that propositions are necessarily equivalent, i.e., $\Box(p \equiv q)$, does not entail that p and q are identical. One may consistently assert that there are propositions p and q such that $\Box(p\equiv q) \,\&\, p\neq q$.

(**195**) **Theorems:** Conditions Implying Distinctness of Relations. For any $n \geq 0$, it is straightforward to show that (.1) relations F^n and G^n that might fail to be materially equivalent are distinct, and (.2) if it is possible that F^n's being such that φ is not equivalent to G^n's being such that φ, then F^n and G^n are distinct:

(.1) $\Diamond\neg\forall x_1 \ldots \forall x_n(F^n x_1 \ldots x_n \equiv G^n x_1 \ldots x_n) \rightarrow F^n \neq G^n$ $(n \geq 0)$

(.2) $\Diamond\neg(\varphi \equiv \varphi') \rightarrow F^n \neq G^n$, whenever G^n is substitutable for F^n in φ and φ' is the result of substituting G^n for one or more free occurrences of F^n in φ. $(n \geq 0)$

As easy corollaries, we have (.3) if F^n and G^n aren't materially equivalent, they are distinct; and (.4) if F^n's being such that φ is not equivalent to G^n's being such that φ, then F^n and G^n are distinct:

(.3) $\neg\forall x_1 \ldots \forall x_n(F^n x_1 \ldots x_n \equiv G^n x_1 \ldots x_n) \rightarrow F^n \neq G^n$

(.4) $\neg(\varphi \equiv \varphi') \rightarrow F^n \neq G^n$, whenever G^n is substitutable for F^n in φ and φ' is the result of substituting G^n for one or more free occurrences of F^n in φ.

(**196**) **Definition:** Definition of Relation Negation. We define $\overline{F}^n$ ('*not-F^n*') as *being $x_1, \ldots, x_n$ such that it is not the case that $F^n x_1 \ldots x_n$* $(n \geq 0)$:

$$\overline{F}^n =_{df} [\lambda x_1 \ldots x_n\, \neg F^n x_1 \ldots x_n] \qquad (n \geq 0)$$

Given the inferential role of definitions, as described in (73), the following theorems holds.

(**197**) **Theorems:** Facts about the Relation Negation Operator. It is axiomatic and so a theorem that (.1) $[\lambda x_1 \ldots x_n\, \neg\Pi^n x_1 \ldots x_n]\downarrow$ for any appropriate n-ary relation term Π^n $(n \geq 0)$. So it follows, for any relation term $\overline{\Pi}$ obtained by applying the negation operator to a relation term Π, that (.2) $\overline{\Pi}$ can be identified with its definiens, thereby implying that (.3) every negated relation term is significant:

(.1) $[\lambda x_1 \ldots x_n\, \neg\Pi^n x_1 \ldots x_n]\downarrow$, where Π^n is any n-ary relation term in which $x_1,\ldots,x_n$ don't occur free

(.2) $\overline{\Pi} = [\lambda x_1 \ldots x_n\, \neg\Pi x_1 \ldots x_n]$, where Π is any n-ary relation term ($n \geq 0$) in which $x_1,\ldots,x_n$ don't occur free.

(.3) $\overline{\Pi}\downarrow$, for any relation term Π

It may be of interest to consider the relation that $[\lambda x_1 \ldots x_n\, \neg\Pi^n x_1 \ldots x_n]$ signifies when $\neg(\Pi\downarrow)$ (assuming $x_1,\ldots,x_n$ don't occur free in Π). In that case, the contrapositive of axiom (39.5.a), which tells us that $\neg(\Pi\downarrow) \to \neg\Pi x_1 \ldots x_n$ ($n \geq 0$), implies $\neg\Pi x_1 \ldots x_n$. This holds for any $x_1,\ldots,x_n$, and so in the case we're considering, the λ-expression $[\lambda x_1 \ldots x_n\, \neg\Pi^n x_1 \ldots x_n]$ signifies a universal n-ary relation, i.e., one exemplified by every sequence of n objects. So, even in the case where Π fails to be significant, $\overline{\Pi}$ signifies what $[\lambda x_1 \ldots x_n\, \neg\Pi^n x_1 \ldots x_n]$ signifies.

(**198**) **Remark:** A Reminder About Definitions by Identity. If the previous theorem comes as a surprise, then it would serve well to review the conventions for, and inferential role of, definitions-by-=, as described in (17), (73), and (120). First, convention (17.2) permits us to use definition (196) instead of the following, less easy-to-read and less user-friendly definition, which is otherwise strictly required:

$$\overline{\Pi}^n =_{df} [\lambda \nu_1 \ldots \nu_n\, \neg\Pi^n \nu_1 \ldots \nu_n], \qquad (n \geq 0)$$
where $\nu_1,\ldots,\nu_n$ are any distinct variables that don't occur free in Π^n

Thus, (196) stipulates that the overline is a relation term-forming operator that applies to every relation term whatsoever.

Second, note that (196) has a single free variable, F^n, in both the definiens and definiendum, and so has the form $\tau(\alpha) =_{df} \sigma(\alpha)$. The Rule of Definition by Identity (73) then guarantees, for any n-ary relation term Π ($n \geq 0$) in which $x_1,\ldots,x_n$ don't occur free, that the following is a necessary axiom:

(ϑ) $([\lambda x_1 \ldots x_n\, \neg\Pi x_1 \ldots x_n]\downarrow \to \overline{\Pi} = [\lambda x_1 \ldots x_n\, \neg\Pi x_1 \ldots x_n])\ \&$
$(\neg[\lambda x_1 \ldots x_n\, \neg\Pi x_1 \ldots x_n]\downarrow \to \neg\overline{\Pi}\downarrow)$

And the derived Rule of Identity by Definition (120.1) implies, when the premise set Γ is empty, that the following holds:

(ζ) If $\vdash [\lambda x_1 \ldots x_n\, \neg\Pi x_1 \ldots x_n]\downarrow$, then $\vdash \overline{\Pi} = [\lambda x_1 \ldots x_n\, \neg\Pi x_1 \ldots x_n]$

Third, theorem (197.1) is that $[\lambda x_1 \ldots x_n\ \neg\Pi x_1 \ldots x_n]\downarrow$, provided Π has no free occurrences of $x_1,\ldots,x_n$. So an identity between $\overline{\Pi}$ and $[\lambda x_1 \ldots x_n\ \neg\Pi x_1 \ldots x_n]$ provably holds, thereby implying by (107.1) that $\overline{\Pi}$ is provably significant, even if Π isn't.

Although $\overline{\Pi}$ can be put into an equation with a significant λ-expression even when Π itself is empty, it seems intuitive to regard $\overline{\Pi}$ as a kind of *impractical*, defined term when Π is empty. There will be other examples of such terms, i.e., defined terms that have a significance even though they are the result of applying an operation to a term that is empty. Though there are ways to further refine the Rule of Definition by Identity (and its consequences) so as to rule out impractical terms (i.e., always ensure that *defined* terms are empty whenever their arguments are empty), such refinements have a cost in system complexity without much gain to show for it. So we'll forego such further refinements in this text. The reader may find a full discussion of the issues involved in Remark (283).

(**199**) **Theorems:** Relations and their Negations. The following consequences of the definition of relation negation are relatively straightforward:

(.1) $\overline{F}^n x_1 \ldots x_n \equiv \neg F^n x_1 \ldots x_n$ $\quad (n \geq 0)$

(.2) $\neg\overline{F}^n x_1 \ldots x_n \equiv F^n x_1 \ldots x_n$ $\quad (n \geq 0)$

(.3) $\overline{p} \equiv \neg p$

(.4) $\neg\overline{p} \equiv p$

(.5) $F^n \neq \overline{F}^n$ $\quad (n \geq 0)$

(.6) $p \neq \overline{p}$

The following are also consequences:

(.7) $\overline{p} = \neg p$

(.8) $p = q \rightarrow \neg p = \neg q$

(.9) $p = q \rightarrow \overline{p} = \overline{q}$

(**200**) **Definitions:** Noncontingent and Contingent Relations. Remembering the reasons why we sometimes have to include existence clauses in the definiens of definitions-by-$\equiv$ (36), we may say: (.1) F^n ($n \geq 0$) is *necessary* just in case necessarily, all objects $x_1,\ldots,x_n$ are such that $x_1,\ldots,x_n$ exemplify F^n; (.2) F^n is *impossible* just in case F^n exists and necessarily, all objects $x_1,\ldots,x_n$ are such that $x_1,\ldots,x_n$ fail to exemplify F^n; (.3) F^n is *noncontingent* whenever it is necessary or impossible; and (.4) F^n is *contingent* whenever it is neither necessary nor impossible:

(.1) $Necessary(F^n) \equiv_{df} \Box\forall x_1 \ldots \forall x_n F^n x_1 \ldots x_n$ $(n \geq 0)$

(.2) $Impossible(F^n) \equiv_{df} F^n\!\downarrow \,\&\, \Box\forall x_1 \ldots \forall x_n \neg F^n x_1 \ldots x_n$ $(n \geq 0)$

(.3) $NonContingent(F^n) \equiv_{df} Necessary(F^n) \vee Impossible(F^n)$ $(n \geq 0)$

(.4) $Contingent(F^n) \equiv_{df} F^n\!\downarrow \,\&\, \neg(Necessary(F^n) \vee Impossible(F^n))$ $(n \geq 0)$

The explicit existence clauses in (.2) and (.4) allow us to prove $\neg Impossible(\Pi^n)$ and $\neg Contingent(\Pi^n)$ when Π^n is provably empty.

(**201**) **Remark:** Observations About Definitions by Equivalence. The following observations serve as reminders for those who may have skipped the discussion of definitions-by-$\equiv$ in Remarks (17.2) and (36). The points made below by way of the definitions (200.1) and (200.2) are completely general.

By Convention (17.2), definition (200.1) and (200.2) are shorthand, respectively, for the following, in which Π^n is any n-ary relation term $(n \geq 0)$:

(.1) $Necessary(\Pi^n) \equiv_{df} \Box\forall x_1 \ldots \forall x_n \Pi^n x_1 \ldots x_n$ $(n \geq 0)$

(.2) $Impossible(\Pi^n) \equiv_{df} \Pi^n\!\downarrow \,\&\, \Box\forall x_1 \ldots \forall x_n \neg \Pi^n x_1 \ldots x_n$ $(n \geq 0)$

Given the inferential role of definitions-by-$\equiv$ described in (72) and (90), (.1) and (.2) (respectively), and (200.1) and (200.2) (respectively), introduce, as new (modally strict) theorems, the instances (and their closures) of the following:

(.1′) $Necessary(\Pi^n) \equiv \Box\forall x_1 \ldots \forall x_n \Pi^n x_1 \ldots x_n$ $(n \geq 0)$

(.2′) $Impossible(\Pi^n) \equiv (\Pi^n\!\downarrow \,\&\, \Box\forall x_1 \ldots \forall x_n \neg \Pi^n x_1 \ldots x_n)$ $(n \geq 0)$

Consider the instances of these theorem schemata when the metavariable Π takes the variable F^n as its value:

(ϑ) $Necessary(F^n) \equiv \Box\forall x_1 \ldots \forall x_n F^n x_1 \ldots x_n$ $(n \geq 0)$

(ξ) $Impossible(F^n) \equiv (F^n\!\downarrow \,\&\, \Box\forall x_1 \ldots \forall x_n \neg F^n x_1 \ldots x_n)$ $(n \geq 0)$

Clearly, these result by substituting $\equiv$ for $\equiv_{df}$ in (200.1) and (200.2). Furthermore, (ξ) and the axiom $F^n\!\downarrow$ imply the following, by the special case of Rule $\equiv$S of Biconditional Simplification (91):

$$Impossible(F^n) \equiv \Box\forall x_1 \ldots \forall x_n \neg F^n x_1 \ldots x_n \qquad (n \geq 0)$$

This kind of reduction is completely general and applies to any formula definition that (a) has a definiens that includes existence clauses, and (b) is instanced to denoting terms. That is, when we consider the instances of the definition in which known denoting terms (i.e., terms known to be significant by proof or by hypothesis) are substituted for the metavariables, we may *reduce* the resulting modally strict biconditionals implied by the definition to modally strict biconditionals in which the existence clauses have been omitted from the right-side condition.

(**202**) **Theorems:** Facts About Noncontingent and Contingent Properties. The following facts about properties F are now derivable:

(.1) $NonContingent(F) \equiv NonContingent(\overline{F})$

(.2) $Contingent(F) \equiv \Diamond\exists xFx \,\&\, \Diamond\exists x\neg Fx$

(.3) $Contingent(F) \equiv Contingent(\overline{F})$

(**203**) **Theorems:** Some Noncontingent Properties. Since $\vdash [\lambda x\, E!x \rightarrow E!x]\downarrow$, let $L =_{df} [\lambda x\, E!x \rightarrow E!x]$ (i.e., *being concrete if concrete*). Then we have:

(.1) $Necessary(L)$

(.2) $Impossible(\overline{L})$

(.3) $NonContingent(L)$

(.4) $NonContingent(\overline{L})$

(.5) $\exists F\exists G(F \neq G \,\&\, NonContingent(F) \,\&\, NonContingent(G))$,
i.e., there are at least two noncontingent properties.

(**204**) **Lemmas:** A Symmetry. Note that our definitions allow us to more easily derive (.1) it is possible that there is something that exemplifies F but which might not have if and only if it is possible that there is something that doesn't exemplify F but might have:

(.1) $\Diamond\exists x(Fx \,\&\, \Diamond\neg Fx) \equiv \Diamond\exists x(\neg Fx \,\&\, \Diamond Fx)$

But from this, it follows that (.2) it is possible that there is something that exemplifies F but which might not have if and only if it is possible that there is something that exemplifies not-F but might not have:

(.2) $\Diamond\exists x(Fx \,\&\, \Diamond\neg Fx) \equiv \Diamond\exists x(\overline{F}x \,\&\, \Diamond\neg\overline{F}x)$

If we think semantically for the moment and take possible worlds as primitive entities, then (.2) tells us:

> There is a world $\boldsymbol{w}$ and an object $\boldsymbol{o}$, such that both (i) $\boldsymbol{o}$ exemplifies F at $\boldsymbol{w}$ and (ii) for some (other) world $\boldsymbol{w}'$, $\boldsymbol{o}$ fails to exemplify F at $\boldsymbol{w}'$
> if and only if
> There is a world $\boldsymbol{w}$ and an object $\boldsymbol{o}$, such that both (i) $\boldsymbol{o}$ exemplifies not-F at $\boldsymbol{w}$ and (ii) for some (other) world $\boldsymbol{w}'$, $\boldsymbol{o}$ fails to exemplify not-F at $\boldsymbol{w}'$.

Thus (.2) yields a modal symmetry between F and its negation $\overline{F}$.

(**205**) **Theorems:** $E!$ and $\overline{E!}$ are Contingent Properties.

(.1) $\Diamond\exists x(E!x \,\&\, \Diamond\neg E!x)$

(.2) $\Diamond\exists x(\neg E!x \,\&\, \Diamond E!x)$

(.3) $\Diamond\exists x E!x$

(.4) $\Diamond\exists x\neg E!x$

(.5) $Contingent(E!)$

(.6) $Contingent(\overline{E!})$

(.7) $\exists F\exists G(Contingent(F) \,\&\, Contingent(G) \,\&\, F \neq G)$,
i.e., there are at least two contingent properties.

If we suppose that $E!x \,\&\, \Diamond\neg E!x$ asserts that x is contingently concrete and that $\neg E!x \,\&\, \Diamond E!x$ asserts that x is contingently nonconcrete, then (.1) asserts the possible existence of contingently concrete objects and (.2) asserts the possible existence of contingently nonconcrete objects. Linsky and Zalta 1994 show that theorems such as (.2), on this understanding, are consistent with *actualism*, though see Williamson 2013, Menzel 2016 and 2020, for more recent discussion.

On an alternative interpretation of our formalism, in which (a) the quantifier $\exists$ asserts only "there is" (and not "there exists") and (b) the term '$E!$' asserts existence rather than concreteness, then (.2) asserts that it is possible that there are contingently nonexistent objects. This is a way of expressing the view known as *possibilism*. See, for example, Plantinga (1974, 21).

(**206**) **Theorems:** Facts About Property Existence. Given the definition of L as $[\lambda x\, E!x \to E!x]$, we have the following general and specific facts about the existence of properties:

(.1) $NonContingent(F) \to \neg\exists G(Contingent(G) \,\&\, G = F)$

(.2) $Contingent(F) \to \neg\exists G(NonContingent(G) \,\&\, G = F)$

(.3) $L \neq \overline{L} \,\&\, L \neq E! \,\&\, L \neq \overline{E!} \,\&\, \overline{L} \neq E! \,\&\, \overline{L} \neq \overline{E!} \,\&\, E! \neq \overline{E!}$,
i.e., L, $\overline{L}$, $E!$, and $\overline{E!}$ are pairwise distinct.

(**207**) **Theorems:** Facts About Noncontingent and Contingent Propositions. If we focus now just on propositions, the following facts can be established:

(.1) $NonContingent(p) \equiv NonContingent(\overline{p})$

(.2) $Contingent(p) \equiv \Diamond p \,\&\, \Diamond\neg p$

(.3) $Contingent(p) \equiv Contingent(\overline{p})$

(**208**) **Theorems:** Some Noncontingent Propositions. Let p_0 abbreviate $\forall x(E!x \rightarrow E!x)$ or, if you prefer, use a definition-by-= to stipulate $p_0 =_{df} \forall x(E!x \rightarrow E!x)$. Then we have:

(.1) $Necessary(p_0)$

(.2) $Impossible(\overline{p_0})$

(.3) $NonContingent(p_0)$

(.4) $NonContingent(\overline{p_0})$

(.5) $\exists p \exists q(NonContingent(p) \,\&\, NonContingent(q) \,\&\, p \neq q)$,
i.e., there are at least two noncontingent propositions.

(**209**) ⋆**Theorem:** Fact About Concrete-But-Not-Actually-Concrete Objects. It is straightforward to establish, by non-modally strict reasoning, that concrete-but-not-actually-concrete objects don't exist:

$\neg\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$

Though it follows by the T$\Diamond$ schema that this fact is possibly true, we can prove the possibility claim by modally strict reasoning, as shown in the next item.

(**210**) **Theorem:** Distinguished Facts About Objects that Aren't Actually Concrete. It is a modally strict theorem that (.1) a concrete-but-not-actually-concrete object doesn't *actually* exist, i.e.,

(.1) $\neg\mathcal{A}\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$

Without much further reasoning, it follows that (.2) possibly, no concrete-but-not-actually-concrete object exists:

(.2) $\Diamond\neg\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$

As we shall see, (.2) leads to a proof of the existence of a contingent proposition. Note that the proofs of (.1) and (.2) are independent of axiom (45.4).

One other important theorem about objects that aren't actually concrete is that (.3) some possibly concrete object isn't actually concrete:

(.3) $\exists x(\Diamond E!x \,\&\, \neg\mathcal{A}E!x)$

(**211**) **Theorems:** Some Contingent Propositions. The formula $\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$ asserts the existence of a concrete-but-not-actually-concrete object. Let us therefore use a definition-by-= to stipulate: $q_0 =_{df} \exists x(E!x \,\&\, \neg\mathcal{A}E!x)$. Then, by Rule $=_{df}$I, the conjunction of axiom (45.4) and theorem (210.2) becomes $\Diamond q_0 \,\&\, \Diamond\neg q_0$. This makes it easy to see that the following claims regarding propositions are derivable:

(.1) $Contingent(q_0)$

(.2) $\exists p Contingent(p)$

(.3) $Contingent(\overline{q_0})$

(.4) $\exists p \exists q(p \neq q \,\&\, Contingent(p) \,\&\, Contingent(q))$,
i.e., there are at least two contingent propositions.

(**212**) **Theorems:** Facts About Proposition Existence. The following general facts about proposition existence are easily derivable:

(.1) $NonContingent(p) \rightarrow \neg\exists q(Contingent(q) \,\&\, q = p)$

(.2) $Contingent(p) \rightarrow \neg\exists q(NonContingent(q) \,\&\, q = p)$

And given the definitions $p_0 =_{df} \forall x(E!x \rightarrow E!x)$ and $q_0 =_{df} \exists x(E!x \,\&\, \neg\mathcal{A}E!x)$, we have the following specific facts about the existence of propositions:

(.3) $p_0 \neq \overline{p_0} \,\&\, p_0 \neq q_0 \,\&\, p_0 \neq \overline{q_0} \,\&\, \overline{p_0} \neq q_0 \,\&\, \overline{p_0} \neq \overline{q_0} \,\&\, q_0 \neq \overline{q_0}$,
i.e., p_0, $\overline{p_0}$, q_0, and $\overline{q_0}$ are pairwise distinct.

(.4) There are at least four propositions.

Note that our proof that there are contingent propositions is constructive, since we showed that q_0 is possibly true and possibly false.

(**213**) **Definitions:** Contingently True and Contingently False Propositions. Let us say that: (.1) a proposition p is *contingently true* just in case p is true and possibly false, and (.2) p is *contingently false* just in case p is false but possibly true:

(.1) $ContingentlyTrue(p) \;\equiv_{df}\; p \,\&\, \Diamond\neg p$

(.2) $ContingentlyFalse(p) \;\equiv_{df}\; \neg p \,\&\, \Diamond p$

Clearly, if one were to assert *axioms* having the form $ContingentlyTrue(p)$ or $ContingentlyFalse(p)$, such axioms would have to be marked with a $\star$ as modally fragile; see the discussion in Remark (70). By the same reasoning, it should be clear that no theorem of the form $ContingentlyTrue(p)$ or $ContingentlyFalse(p)$ can be be modally strict, on pain of contradiction. For suppose it were provable by modally strict means that *ContingentlyTrue*(p), i.e., $p \,\&\, \Diamond\neg p$. Then it would follow that $\Box(p \,\&\, \Diamond\neg p)$, by RN. But the conjuncts of a necessary conjunction are themselves necessary, in which case it would follow that $\Box p$ and $\Box\Diamond\neg p$. By the T schema, the latter implies $\Diamond\neg p$, which in turn implies $\neg\Box p$, leaving us with a contradiction. And, by analogous reasoning, no theorem of the form *ContingentlyFalse*(p) can be modally strict. Thus, in (215.1)★ and (215.2)★ below, when we derive, respectively, a specific, contingently false proposition and a specific, contingently true proposition, the theorems are not modally strict.

(**214**) **Theorems:** Contingently True (False) vs. Contingent. Our definitions and theorems thus far imply the following. (.1) if p is contingently true, it is contingent; (.2) if p is contingently false, it is contingent; (.3) p is contingently true if and only if not-p is contingently false; and (.4) p is contingently false if and only if not-p is contingently true:

(.1) $\mathit{ContingentlyTrue}(p) \rightarrow \mathit{Contingent}(p)$

(.2) $\mathit{ContingentlyFalse}(p) \rightarrow \mathit{Contingent}(p)$

(.3) $\mathit{ContingentlyTrue}(p) \equiv \mathit{ContingentlyFalse}(\overline{p})$

(.4) $\mathit{ContingentlyFalse}(p) \equiv \mathit{ContingentlyTrue}(\overline{p})$

Finally, two other theorems are worthy of mention. (.5) if p is contingently true and q is necessary, then p is not identical to q; and (.6) if p is contingently false and q is impossible, then p is not identical to q:

(.5) $(\mathit{ContingentlyTrue}(p) \,\&\, \mathit{Necessary}(q)) \rightarrow p \neq q$

(.6) $(\mathit{ContingentlyFalse}(p) \,\&\, \mathit{Impossible}(q)) \rightarrow p \neq q$

Recall that *Necessary*(q) was defined, by the 0-ary cases of (200.1), as $\Box q$, and that *Impossible*(q), by (200.2), the axiom $q\downarrow$ and Rule $\equiv$S (91), is equivalent to $\Box\neg q$.

(**215**) ★**Theorems:** A Contingently False Proposition and a Contingently True One. If we continue to let the constant q_0 be defined as $\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$, it follows that (.1) q_0 is contingently false, and (.2) not-q_0 is contingently true:

(.1) $\mathit{ContingentlyFalse}(q_0)$

(.2) $\mathit{ContingentlyTrue}(\overline{q_0})$

As to be expected from the discussion in (213), the reasoning in each case is not modally strict.

(**216**) **Remark:** Actually Contingently True and False Propositions. Note that if one were to define:

$$ActuallyContingentlyTrue(p) \equiv_{df} \mathcal{A}p \,\&\, \Diamond\neg p$$

$$ActuallyContingentlyFalse(p) \equiv_{df} \mathcal{A}\neg p \,\&\, \Diamond p$$

then where q_0 is again defined as $\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$, one can immediately establish, as modally strict theorems, that:

(ϑ_1) $ActuallyContingentlyTrue(p) \equiv ActuallyContingentlyFalse(\overline{p})$

(ϑ_2) $ActuallyContingentlyFalse(p) \equiv ActuallyContingentlyTrue(\overline{p})$

(ξ) $ActuallyContingentlyFalse(q_0)$

(ζ) $ActuallyContingentlyTrue(\overline{q_0})$

We give the proof of (ϑ_1) in a footnote but leave the proof of (ϑ_2) to the reader.[168] Here is a proof of (ξ):

> From (210.1), it follows by axiom (44.1) that $\mathcal{A}\neg\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$. So, by definition of q_0, $\mathcal{A}\neg q_0$. Independently by axiom (45.4), we know $\Diamond q_0$. Hence $ActuallyContingentlyFalse(q_0)$.

We leave the proof of (ζ) to the reader.

(**217**) **Theorems:** There Are Contingently True Propositions and Contingently False Ones. If we were to immediately infer $\exists p ContingentlyFalse(p)$ from theorem (215.1)★, and immediately infer $\exists p ContingentlyTrue(p)$ from (215.2)★, the resulting theorems would fail to be modally strict. But, in fact, we can derive these theorems by modally strict means:

(.1) $\exists p ContingentlyTrue(p)$

(.2) $\exists p ContingentlyFalse(p)$

Thus, we need not appeal to (215.2)★ to derive (.1) or appeal to (215.1)★ to derive (.2).

(**218**) **Remark:** On the Existence of Contingently True Propositions and Contingently False Ones. Many philosophers would grant (217.1) and (217.2), i.e., that there are contingently true and contingently false propositions. But by deriving these as theorems, our theory answers a philosophical question, namely, what resources are needed to (a) *prove* rather than assert the existence of a contingently true proposition (and a contingently false proposition) and (b) do

[168] We prove both directions: ($\rightarrow$) Assume $ActuallyContingentlyTrue(p)$. Then $\mathcal{A}p \,\&\, \Diamond\neg p$. We want to show $\mathcal{A}\neg\overline{p} \,\&\, \Diamond\overline{p}$. But by (199.4), we know that $\neg\overline{p} \equiv p$. So by a Rule of Substitution, $\mathcal{A}p$ implies $\mathcal{A}\neg\overline{p}$. Analogously, by (199.3), we know that $\overline{p} \equiv \neg p$. So by a Rule of Substitution, $\Diamond\neg p$ implies $\Diamond\overline{p}$. ($\leftarrow$) By analogous reasoning.

so without asserting, for some particular individual x and property F, either $Fx \,\&\, \Diamond\neg Fx$ or $\neg Fx \,\&\, \Diamond Fx$?

The present theory gives us both modally strict and non-modally strict proofs that such propositions exist, and axiom (45.4) plays a role in these arguments. This is another virtue of taking (45.4) as an axiom. The non-modally strict proofs are immediate existential generalizations of (215.1)★ and (215.2)★ and, as such, are constructive (specific witnesses to the existential claims can be identified). The modally strict proofs of (217.1) and (217.2) aren't constructive and, indeed, can't be; as we've seen, there can't be a modally strict argument for the contingent truth or falsehood of a particular proposition that would serve as a witness.

Thus, we sometimes have a choice of strategies when arguing for claims that depend on the existence of contingently true or contingently false propositions. If one directly uses (215.1)★ or (215.2)★ in the reasoning for the claim, the resulting theorem is not modally strict. However, in some cases, it may be that the very same theorem can be proved by modally strict means from (217.1) or (217.2). Here it is important to recall and extend our discussion in Remark (70). That discussion explained (a) how there can be modally strict derivations of contingent conclusions φ from contingent premises Γ, and (b) how we can still use RN in such cases to conclude $\Box\Gamma \vdash \Box\varphi$. In the case of (217.1) and (217.2), we proved a modally strict theorem, not just deriving a conclusion from some premises. To prove the theorem, we have to *discharge* any premises used in the proof and, in particular, discharge any contingent premises used in the proof. This guarantees that the theorem rests on no contingent assumptions. For example, one may invoke the modally strict theorem that there are contingently true propositions (217.1) and then assume that some arbitrarily chosen proposition is a witness. Then we can we reason from this assumption by modally strict means and once we discharge the assumption by ∃E, the conclusions we reach depend only on the modally strict fact that there are contingent propositions, and so these conclusions are modally strict.

To be maximally explicit, let's rehearse the 'Simpler Proof' for (217.2) given in the Appendix. The simpler proof goes as follows:

> By (217.1), we know $\exists p\,\mathit{ContingentlyTrue}(p)$. Let r be such a proposition, so that we know $\mathit{ContingentlyTrue}(r)$. Then since r exists by hypothesis, we may infer $\mathit{ContingentlyFalse}(\overline{r})$, by (214.3). But $\overline{r}\downarrow$, and so by ∃I, $\exists p\,\mathit{ContingentlyFalse}(p)$. Hence, by ∃E, $\exists p\,\mathit{ContingentlyFalse}(p)$. ⋈

In this reasoning, the assumption $\mathit{ContingentlyTrue}(r)$ is *discharged*, and since the theorem depends only on (217.1), it is modally strict.

Exercise. Consider argument (A), which attempts to derive a contradiction in our system, and before reading the discussion that follows it, explain where it goes wrong:

(A) By (217.1), we know $\exists p ContingentlyTrue(p)$. Suppose p_1 is an arbitrary such proposition, so that we know $ContingentlyTrue(p_1)$. Then, by definition (213.1), $p_1 \mathbin{\&} \Diamond\neg p_1$. So by &E, p_1. Since this reasoning is modally strict, we may use RN to infer $\Box p_1$, i.e., by (158.12), $\neg\Diamond\neg p_1$. But this contradicts $\Diamond\neg p_1$.

Argument (A) is interesting because it demonstrates a subtle misuse of RN. In argument (A), it is assumed that p_1 is a contingently true proposition. This assumption implies, by modally strict reasoning, that p_1 is true. But RN can't be applied to conclude that $\Box p_1$, since p_1 isn't a theorem but was instead derived from an assumption (all we're entitled to conclude by RN at this point is that there is a derivation of $\Box p_1$ from $\Box ContingentlyTrue(p_1)$). This doesn't get us into trouble since we can't prove the latter. So the reasoning in (A) isn't valid. One may not use RN to necessitate a conclusion derived from premises, unless those premises are known to be necessary.

(**219**) **Theorems:** There are Contingently Exemplified and Contingently Empty Properties. By making use of contingent propositions, we can go further than theorem (205.7), which tells us that there are at least two contingent properties; we can now establish that (.1) there are contingently exemplified properties, i.e., properties F such that some object exemplifies F but might not have, and (.2) there are contingently unexemplified properties (i.e., properties F such that some object that doesn't exemplify F might have):

(.1) $\exists F \exists x(Fx \mathbin{\&} \Diamond\neg Fx)$

(.2) $\exists F \exists x(\neg Fx \mathbin{\&} \Diamond Fx)$

In the Appendix, the modally strict reasoning for (.1) starts with theorem (217.1), i.e., that $\exists p ContingentlyTrue(p)$ and proceeds by assuming that p_1 is an arbitrarily chosen such proposition, i.e., that $p_1 \mathbin{\&} \Diamond\neg p_1$. We then show that the property $[\lambda x\, p_1]$ is a witness that establishes the theorem, i.e., that any fixed, but arbitrary object y exemplifies $[\lambda x\, p_1]$ but doesn't necessarily exemplify $[\lambda x p_1]$.[169] The proof illustrates the discussions in Remarks (70) and (218); the temporary assumption $p_1 \mathbin{\&} \Diamond\neg p_1$ isn't a necessary truth, but since it constitutes an arbitrary instance of a modally strict, and hence necessarily true,

[169] Propositional properties like $[\lambda x\, p_1]$ will be discussed at some length in Section 9.12. For now, all one needs to know that they are perfectly well-behaved – every λ-expression $[\lambda x\, \varphi]$ in which x doesn't occur free in φ is significant, by (39.2), and so all such expressions are governed by β-Conversion. Thus, by β-Conversion, $[\lambda x\, p_1]y \equiv p_1$ holds generally.

If one prefers, (.1) can be proved by appeal to a non-propositional property. For example, $[\lambda x\, (E!x \rightarrow E!x) \mathbin{\&} p_1]$ would do the job. The variable x in the expression denoting this property isn't vacuously bound by the λ. But the property will suffice since the proof requires an appeal to a universal property (i.e., everything exemplifies it), but one that is only contingently a universal property.

existence claim, it can be used and discharged in a way that doesn't subvert the modally strict reasoning in the proof.

There is a non-modally strict argument for (.1), namely, let Q in the proof be $[\lambda x\, \overline{q_0}]$, where q_0 is defined as $\exists x(E!x \,\&\, \neg \mathcal{A}E!x)$. By (215.2)★, we know $\overline{q_0}$ is contingently true, i.e., such that $\overline{q_0} \,\&\, \Diamond\neg\overline{q_0}$. By reasoning analogous to that used in the proof of (.1) in the Appendix, one can show that $[\lambda x\, \overline{q_0}]$ is exemplified but not possibly not exemplified. Intuitively, since $\Diamond\neg\overline{q_0}$, there is a possible world where $\overline{q_0}$ is false and, at that world, objects fail to exemplify $[\lambda x \overline{q_0}]$. So it would be a witness to the claim that some property is contingently exemplified, though the reasoning wouldn't be modally strict.

(220) **Remark:** Equivalent But Possibly Not Equivalent Properties. It is straightforward to *apply* object theory by adding theoretical properties to the system and then asserting, as an axiom, that certain theoretical properties are materially equivalent but not necessarily equivalent. A well-known example of this phenomena is due to Quine (1951, 21–2): Let H be the property *being a creature with a heart* and let K be the property *being a creature with a kidney*. Then let us suppose, for the sake of the example, that $\forall x(Hx \equiv Kx)$ represents a biological fact and is asserted as an axiom, but not a necessary axiom. After all, it seems doubtful that this claim is necessary; from a modal point of view, it seems reasonable to assert, also as an axiom, that it is possible that there be creatures with a heart but without kidneys or creatures with kidneys but without hearts. While such extensions of the theory are straightforward, a more interesting question arises, namely, can we prove *without* adding assumptions such as the foregoing that there are properties which are materially equivalent but not necessarily equivalent? The following theorems establish that we can. Indeed, these theorems can be strengthened to show that for every property F, there is a property G that is materially equivalent to F but not necessarily equivalent to F.

(221) **Theorem:** Equivalent But Not Necessarily Equivalent Properties. (.1) There are properties F and G such that F and G are materially equivalent but not necessarily equivalent:

(.1) $\exists F\exists G(\forall x(Fx \equiv Gx) \,\&\, \Diamond\neg\forall x(Fx \equiv Gx))$

In the Appendix, the modally strict reasoning for this claim again starts with theorem (217.1), i.e., that $\exists p(\mathit{ContingentlyTrue}(p)$ and proceeds by assuming that p_1 is an arbitrarily chosen such proposition, i.e., that $p_1 \,\&\, \Diamond\neg p_1$. We then show that the properties $[\lambda x\, E!x \rightarrow E!x]$ (= L) and $[\lambda x\, p_1]$ (= Q) are witnesses to our theorem, i.e., that L and Q are materially equivalent but not necessarily equivalent. The proof is another illustration of the discussions in Remarks (70) and (218) about the circumstances in which assumptions that aren't necessarily true (e.g., our assumption that $p_1 \,\&\, \Diamond\neg p_1$) can be used temporarily and discharged without subverting modally strict reasoning.

By analogous reasoning from the existence of contingently false propositions, it can be shown that (.2) there are properties F and G that are not materially equivalent but possibly materially equivalent:

(.2) $\exists F \exists G(\neg\forall x(Fx \equiv Gx) \,\&\, \Diamond\forall x(Fx \equiv Gx))$

Finally, the following fact proves useful, namely (.3) there are properties F and G that are actually not equivalent but possibly equivalent:

(.3) $\exists F \exists G(\mathcal{A}\neg\forall x(Fx \equiv Gx) \,\&\, \Diamond\forall x(Fx \equiv Gx))$

In fact, even stronger claims follow, namely: (.4) for every property F, there is a property G that is equivalent to F but possibly not equivalent; (.5) for every property F, there is a property G that is not equivalent to F but possibly so; and (.6) for every property F, there is a property G that is actually not equivalent to F but which is possibly equivalent to F:

(.4) $\forall F \exists G(\forall x(Fx \equiv Gx) \,\&\, \Diamond\neg\forall x(Fx \equiv Gx))$

(.5) $\forall F \exists G(\neg\forall x(Fx \equiv Gx) \,\&\, \Diamond\forall x(Fx \equiv Gx))$

(.6) $\forall F \exists G(\mathcal{A}\neg\forall x(Fx \equiv Gx) \,\&\, \Diamond\forall x(Fx \equiv Gx))$

While (.4) – (.6) don't play an immediate role in the theorems that follows, they will prove to be important in a later chapter.

(222) **Theorems:** $O!$ and $A!$ Are Contingent. Recall that a contingent property is defined to be one that is neither necessary nor impossible (200), from which it follows that a property F is contingent just in case it is both possible that there is something that exemplifies F and possible that there is something that fails to exemplify F (202.2). We now show that the properties *being ordinary* and *being abstract* are distinct, contradictory contingent properties:

(.1) $O! \neq A!$

(.2) $O!x \equiv \neg A!x$

(.3) $A!x \equiv \neg O!x$

(.4) $\mathit{Contingent}(O!)$

(.5) $\mathit{Contingent}(A!)$

Moreover, the negations of *being ordinary* and *being abstract* are distinct and contradictory contingent properties:

(.6) $\overline{O!} \neq \overline{A!}$

(.7) $\overline{O!}x \equiv \neg\overline{A!}x$

(.8) $Contingent(\overline{O!})$

(.9) $Contingent(\overline{A!})$

(223) **Definition:** Weakly Contingent Properties. We say that a property F is *weakly contingent* just in case F is contingent but anything that possibly exemplifies F necessarily exemplifies F:

$$WeaklyContingent(F) \ \equiv_{df} \ Contingent(F) \,\&\, \forall x(\Diamond Fx \rightarrow \Box Fx)$$

No existence clause is needed in the definiens (cf. (17.2)), since one is already implied by the definition of $Contingent(F)$ (200.4).

(224) **Theorems:** Facts About Weakly Contingent Properties. (.1) F is weakly contingent iff $\overline{F}$ is weakly contingent; (.2) if F is weakly contingent and G is not, then F is not G:

(.1) $WeaklyContingent(F) \equiv WeaklyContingent(\overline{F})$

(.2) $(WeaklyContingent(F) \,\&\, \neg WeaklyContingent(G)) \rightarrow F \neq G$

(225) **Theorems:** Facts About $O!$, $A!$, $E!$, and L. Using the definition $L =_{df} [\lambda x\, E!x \rightarrow E!x]$, we have the following facts: (.1) *being ordinary* is weakly contingent; (.2) *being abstract* is weakly contingent; (.3) *being concrete* is not weakly contingent; (.4) *being concrete if concrete* is not weakly contingent; (.5) *being ordinary* is distinct from: $E!$, $\overline{E!}$, L, and $\overline{L}$; and (.6) *being abstract* is distinct from: $E!$, $\overline{E!}$, L, and $\overline{L}$:

(.1) $WeaklyContingent(O!)$

(.2) $WeaklyContingent(A!)$

(.3) $\neg WeaklyContingent(E!)$

(.4) $\neg WeaklyContingent(L)$

(.5) $O! \neq E! \,\&\, O! \neq \overline{E!} \,\&\, O! \neq L \,\&\, O! \neq \overline{L}$

(.6) $A! \neq E! \,\&\, A! \neq \overline{E!} \,\&\, A! \neq L \,\&\, A! \neq \overline{L}$

Note that (.5) and (.6) don't imply that $O! \neq \overline{A!}$ or that $A! \neq \overline{O!}$.

(226) **Theorem:** There Are At Least 16 Properties.[170] It is provable that there are at least sixteen properties, i.e.,

$$\exists F_1 \ldots \exists F_{16}(F_1 \neq F_2 \,\&\, F_1 \neq F_3 \,\&\, \ldots \,\&\, F_1 \neq F_{16} \,\&\, F_2 \neq F_3 \,\&\, \ldots \,\&\, F_2 \neq F_{16} \,\&\, \ldots \,\&\, F_{15} \neq F_{16})$$

Exercise: Show how this theorem could be used to prove that there are at least 65,536 (i.e., 2^{16}) abstract objects.

[170] Contributed by Daniel Kirchner, personal communication, 29 May 2018.

9.11 The Theory of Objects

(227) **Theorem:** Ordinary vs. Abstract Objects. Our system implies that (.1) necessarily, there exists an object that exemplifies *being ordinary*; (.2) necessarily, there exists an object that exemplifies *being abstract*; (.3) necessarily, it is not the case that every object exemplifies *being ordinary*; (.4) necessarily, it is not the case that every object exemplifies *being abstract*; and (.5) necessarily, it is not the case that every object exemplifies *being concrete*:

(.1) $\Box\exists x O!x$

(.2) $\Box\exists x A!x$

(.3) $\Box\neg\forall x O!x$

(.4) $\Box\neg\forall x A!x$

(.5) $\Box\neg\forall x E!x$

It is important to note that none of these imply $\exists x E!x$, i.e., these theorems don't imply the existence of any concrete objects. The existence of concrete objects is an empirical matter, subject to *a posteriori* investigation rather than a derivation from axioms asserted *a priori*.

(228) **Theorems:** The Domain of Objects is Partitioned. Since it follows by GEN from theorem (115.5) that $\forall x(O!x \vee A!x)$, we can show that the domain of objects is partitioned if we now establish that no object is both ordinary and abstract:

$$\neg\exists x(O!x \,\&\, A!x)$$

This is a modally-strict theorem and, hence, a necessary truth.

9.11.1 Ordinary Objects

We've seen that there is no general relation of identity, i.e., no binary relation F such that $\forall x\forall y(Fxy \equiv x=y)$ (192.5). It follows that the expression $[\lambda xy\, x=y]$ isn't significant. But a relation of identity can be defined with respect to the ordinary objects. As it turns out, by simply restricting the defined condition $x=y$ to ordinary objects, we can prove the existence of a relation that holds between objects x and y just in case they are identical ordinary objects. The principles governing this relation, and governing ordinary objects generally, we be investigated in this section.

(229) **Theorem:** A Distinguished Relation. Using axiom (49), we may prove that *being ordinary objects that are identical* exists:

$$[\lambda xy\, O!x \,\&\, O!y \,\&\, x = y]\downarrow$$

We may therefore use this relation to introduce a relation of identity with respect to ordinary objects.

(**230**) **Definition:** The Identity$_E$ Relation. We define the binary relation *being identical ordinary objects* (or *being identical*$_E$) as: being an individual x and an individual y such that x exemplifies being ordinary, y exemplifies being ordinary, and x is identical to y:

$$=_E \;=_{df}\; [\lambda xy\, O!x \,\&\, O!y \,\&\, x = y]$$

Since the definiens is significant, our theory of definitions guarantees that $=_E\downarrow$. The Rule of Identity by Definition (120.1) yields the claim $=_E \;=\; [\lambda xy\, O!x \,\&\, O!y \,\&\, x = y]$ as a theorem. So by theorem (107.1), it follows that $=_E\downarrow$.

(**231**) **Rewrite Convention:** Identity$_E$ Infix Notation. We adopt, as a convention, the following infix notation for formulas involving the new binary relation symbol $=_E$. Where κ_1 and κ_2 are any two individual terms, we henceforth write $\kappa_1 =_E \kappa_2$ instead of $=_E \kappa_1\kappa_2$. So where κ_1 is the variable x and κ_2 is the constant a, we write $x =_E a$ instead of $=_E xa$. When κ_1 is a and κ_2 is $\imath x\varphi$ (φ any formula), then we write $a =_E \imath x\varphi$ instead of $=_E a\imath x\varphi$. And we write $\imath x\varphi =_E \imath z\psi$ instead of $=_E \imath x\varphi\imath z\psi$.

It is important to note here that if either κ_1 or κ_2 is a description that fails to have a denotation, i.e., if either $\neg\kappa_1\downarrow$ or $\neg\kappa_2\downarrow$, then $\neg(\kappa_1 =_E \kappa_2)$ will be derivable. To see why, we note that $\kappa_1 =_E \kappa_2 \to \kappa_1\downarrow$ and $\kappa_1 =_E \kappa_2 \to \kappa_2\downarrow$ will both be implied by axiom (39.5.a), since κ_1 and κ_2 are primary terms of $\kappa_1 =_E \kappa_2$. So if either $\neg\kappa_1\downarrow$ or $\neg\kappa_2\downarrow$, then $\neg(\kappa_1 =_E \kappa_2)$ will be derivable by Modus Tollens.[171]

The reader is encouraged to review explanatory Remark (232) below, on nested λ-expressions and layers of definition. That Remark contains an important preview of the methods that will become available for unpacking the defined notation in a λ-expression such as $[\lambda x\, x =_E a]$.

The formula $x =_E y$ is therefore unlike, and is to be rigorously distinguished from, the formula $x = y$. The expressions $[\lambda xy\, x =_E y]$, $[\lambda x\; x =_E y]$, $[\lambda x\; a =_E y]$, $[\lambda y\; x =_E y]$, and $[\lambda y\; a =_E y]$ are provably significant; they are all instances of (39.2), even if a were a defined constant instead of primitive, since no variable bound by the λ occurs in encoding position in the matrix. By contrast, $[\lambda xy\, x = y]$ is not a core λ-expression, for the reasons described in the discussions that accompany (17.3) and (23.1).

[171] It is worth remembering here that $\tau\downarrow$ doesn't imply $\sigma\downarrow$ when σ is a proper subterm of τ. So axioms (39.5.a) and (39.5.b) don't imply that *every* subterm in a true exemplification or encoding formula has a denotation; only the primary terms. Recall the example from the discussion in Remark (40): axiom (39.2) asserts that $[\lambda x\, Rx\imath z(Pz \,\&\, \neg Pz)]$ is significant, even though its subterm $\imath z(Pz \,\&\, \neg Pz)$ will be provably empty. (It is provable that nothing exemplifies the property signified by the λ-expression.) See Remark (155) for further examples.

(232) **Remark:** Digression on Nested λ-expressions and Layers of Definition. Given the foregoing discussion, it is important to note how the layers of definition are compounding and how some λ-expressions contain nested occurrences of other λ-expressions. For example, by the convention for infix notation (231), the λ-expression:

$$[\lambda x\, x =_E b]$$

is shorthand for:

$$[\lambda x\ =_E xb]$$

But the definition of the binary relation constant $=_E$ (230) will imply the identity claim $=_E = [\lambda xy\, O!x \,\&\, O!y \,\&\, x = y]$, and since the substitution of identicals holds universally, we can replace the $=_E$ by its definiens in any context. So the last λ-expression displayed above, $[\lambda x\ =_E xb]$, can be expanded, in any context in which it appears, to:

$$[\lambda x\, [\lambda xy\, O!x \,\&\, O!y \,\&\, x = y]xb]$$

Now definition (22.1) of the unary relation constant $O!$ will imply the identity claim $O! = [\lambda x\, \Diamond E!x]$ and so we can replace $O!$ by $[\lambda x\, \Diamond E!x]$ in any context. So we can, in turn, replace the last λ-expression displayed above, in any context, by:

$$(\vartheta)\quad [\lambda x\, [\lambda xy\, [\lambda x\, \Diamond E!x]x \,\&\, [\lambda x\, \Diamond E!x]y \,\&\, x = y]xb]$$

Now at this point, (ϑ) still contains the defined symbols $\Diamond$, &, and =, but we cannot further expand the λ-expression by applying the definitions of these symbols. These are syncategorematic symbols introduced by definitions-by-$\equiv$. The Rule of Substitution derived in (160.3) allows us to substitute the definiens for the definiendum in a definition-by-$\equiv$ only when the definiendum occurs as a *subformula* within some formula; the formulas containing $\Diamond$, &, and = in (ϑ) do not occur as subformulas of (ϑ). Indeed, *subformula of* is not defined for terms generally and so not defined for n-ary λ-expressions ($n \geq 1$). So we can't further unpack (ϑ).

However, asserting that an individual, say a, exemplifies (ϑ), can be simplified by the principle of β-Conversion (48.2). That is, if we assert:

$$[\lambda x\, [\lambda xy\, [\lambda x\, \Diamond E!x]x \,\&\, [\lambda x\, \Diamond E!x]y \,\&\, x = y]xb]a$$

then since (ϑ) is a significant λ-expression, β-Conversion will allow us to infer:

$$[\lambda xy\, [\lambda x\, \Diamond E!x]x \,\&\, [\lambda x\, \Diamond E!x]y \,\&\, x = y]ab$$

Then by a further application of β-Conversion, we can infer:

$$[\lambda x\, \Diamond E!x]a \,\&\, [\lambda x\, \Diamond E!x]b \,\&\, a = b$$

And our deductive system will then allow us to infer:

$$\Diamond E!a \;\&\; \Diamond E!b \;\&\; a = b$$

We leave further analysis to the reader; classical theorems and rules of inference will allow one to apply the definitions of &, $\Diamond$, and = to the subformulas of the above, to expose the primitive notions deeply nested inside the matrix of $[\lambda x\, x =_E b]$.

(**233**) **Theorems:** Useful Theorems About Identity$_E$ and Identity. The following are simple, but useful theorems: (.1) x and y are identical$_E$ if and only if x exemplifies being ordinary, y exemplifies being ordinary, and x and y are identical; (.2) whenever objects are identical$_E$, they are identical; and (.3) x and y are identical$_E$ if and only if x exemplifies being ordinary, y exemplifies being ordinary, and x and y necessarily exemplify the same properties:

(.1) $x =_E y \equiv (O!x \;\&\; O!y \;\&\; x = y)$

(.2) $x =_E y \rightarrow x = y$

(.3) $x =_E y \equiv (O!x \;\&\; O!y \;\&\; \Box\forall F(Fx \equiv Fy))$

(**234**) **Theorems:** Identity$_E$ is Modally Collapsed. (.1) objects x and y are identical$_E$ if and only if they are necessarily identical$_E$; (.2) objects x and y are possibly identical$_E$ if and only if they are identical$_E$; and (.3) x and y are possibly identical$_E$ if and only if they are necessarily identical$_E$:

(.1) $x =_E y \equiv \Box x =_E y$

(.2) $\Diamond x =_E y \equiv x =_E y$

(.3) $\Diamond x =_E y \equiv \Box x =_E y$

Thus, (.3) establishes that the formula $x =_E y$ is modally collapsed and that identity$_E$ is indifferent to modal distinctions. We'll see further consequences of this below. Note that in Chapter 12, when we define *rigid* relations (571.1), it will follow from (.1) above that the relation $=_E$ is a rigid relation.

(**235**) **Rewrite Conventions:** Distinctness$_E$. Note that since $=_E$ is a (defined) relation term (230) and $^-$ is an operation on relation terms, $\overline{=_E}$ is a well-defined symbol of our language. We adopt the following conventions for the defined relation symbol $\overline{=_E}$:

(.1) We henceforth write $\neq_E$ instead of $\overline{=_E}$.

(.2) Where κ_1 and κ_2 are any individual terms, we henceforth write $\kappa_1 \neq_E \kappa_2$ instead of $\neq_E \kappa_1 \kappa_2$.

For example, these conventions not only allow us to write $x \neq_E y$ instead of $\overline{=_E}xy$, but also $[\lambda xy\, x \neq_E y]$ instead of $[\lambda xy\, \overline{=_E}xy]$.

(**236**) **Theorem:** Equivalence of Distinctness$_E$ and Not Identical$_E$.

$$x \neq_E y \equiv \neg(x =_E y)$$

It may come as a surprise that the proof of this theorem is not immediate but rather requires an appeal to (39.2), (48.2), (196), (231), (235.1), and (235.2).

Though this theorem holds generally for all objects x and y, it may fail for arbitrary individual terms κ and κ': the schemas $\kappa \neq_E \kappa'$ and $\neg(\kappa =_E \kappa')$ are *not* equivalent. For if either $\neg\kappa\downarrow$ or $\neg\kappa'\downarrow$ is provable, then both $\kappa =_E \kappa'$ and $\kappa \neq_E \kappa'$ are provably false (they are both exemplification formulas with an empty primary term), in which case $\neg(\kappa =_E \kappa')$ becomes provably true.

(**237**) **Theorems:** Distinctness$_E$ is Modally Collapsed. (.1) objects x and y are distinct$_E$ if and only if necessarily they are distinct$_E$; (.2) objects x and y are possibly distinct$_E$ if and only if they are distinct$_E$; and (.3) objects x and y are possibly distinct$_E$ if and only if they are necessarily distinct$_E$:

(.1) $x \neq_E y \equiv \Box x \neq_E y$

(.2) $\Diamond x \neq_E y \equiv x \neq_E y$

(.3) $\Diamond x \neq_E y \equiv \Box x \neq_E y$

(**238**) **Theorems:** Identity$_E$, Distinctness$_E$, and Actuality.

(.1) $x =_E y \equiv \mathcal{A} x =_E y$

(.2) $x \neq_E y \equiv \mathcal{A} x \neq_E y$

It is important to observe here that these are both modally-strict theorems having the form $\varphi \equiv \mathcal{A}\varphi$. So each constitutes another special group of formulas which, when commuted, have the same form as the modally fragile theorem schema (130.2)$\star$ but which are provable without appealing to that theorem.

(**239**) **Theorems:** Identity$_E$ is an Equivalence Relation on Ordinary Objects.

(.1) $O!x \rightarrow x =_E x$

(.2) $x =_E y \rightarrow y =_E x$

(.3) $(x =_E y \,\&\, y =_E z) \rightarrow x =_E z$

(**240**) **Theorems:** Ordinary Objects and Identity. It now follows that (.1) if either x or y is ordinary, then necessarily, x is identical to y if and only if x is E-identical to y; and (.2) if y is ordinary, then *being identical to* y exists:

(.1) $(O!x \vee O!y) \rightarrow \Box(x = y \equiv x =_E y)$

(.2) $O!y \rightarrow [\lambda x\, x = y]\downarrow$

Note that (.2) tells us that if y is an ordinary object, its haecceity exists. Compare theorem (192.3), which told us that there isn't such a property for *every* object whatsoever.

(241) **Theorem:** Exemplification Indiscernibility Implies Necessary Indiscernibility. It is an interesting fact that if *any* objects x and y are indiscernible, i.e., they exemplify the same properties, then they necessarily do so:

$$\forall F(Fx \equiv Fy) \rightarrow \Box\forall F(Fx \equiv Fy)$$

(242) **Theorem:** Ordinary Objects Obey Leibniz's Law. The identity$_E$ and identity of indiscernible ordinary objects are theorems:

(.1) $(O!x \vee O!y) \rightarrow (\forall F(Fx \equiv Fy) \rightarrow x =_E y)$

(.2) $(O!x \vee O!y) \rightarrow (\forall F(Fx \equiv Fy) \rightarrow x = y)$

These imply that to establish ordinary objects are identical, it suffices to show that they are indiscernible; we don't have to show that they are necessarily indiscernible. So if ordinary objects x, y are distinct$_E$ or distinct *simpliciter*, then we know that there exists a property that distinguishes them.

(243) **Theorem:** Distinct Ordinary Objects Have Distinct Haecceities. It is now relatively straightforward to show that (.1) if x and y are ordinary, then they are distinct iff *being E-identical to* x is distinct from *being E-identical to* y, and (.2) ordinary objects are distinct iff they have distinct haecceities:

(.1) $(O!x \,\&\, O!y) \rightarrow (x \neq y \equiv [\lambda z\, z =_E x] \neq [\lambda z\, z =_E y])$

(.2) $(O!x \,\&\, O!y) \rightarrow (x \neq y \equiv [\lambda z\, z = x] \neq [\lambda z\, z = y])$

(244) **Theorem:** Ordinary Objects Necessarily Fail to Encode Properties.

$$O!x \rightarrow \Box\neg\exists F xF$$

9.11.2 Abstract Objects

(245) **Theorem:** Abstract Objects Obey the Encoding Form of Leibniz's Law. We begin with the fact that (.1) if x and y encode the same properties, then it is necessary that they encode the same properties:

(.1) $\forall F(xF \equiv yF) \rightarrow \Box\forall F(xF \equiv yF)$

A variant of Leibniz's Law now becomes derivable, namely, (.2) whenever abstract objects x, y encode the same properties, they are identical:

(.2) $(A!x \,\&\, A!y) \to (\forall F(xF \equiv yF) \to x = y)$

Thus, to show abstract objects x and y are identical, it suffices to prove that they encode the same properties; we don't have to show that they necessarily encode the same properties. Similarly, it is provable that (.3) whenever an object encodes a property the other fails to encode, they are distinct:

(.3) $(\exists F(xF \,\&\, \neg yF) \vee \exists F(yF \,\&\, \neg xF)) \to x \neq y)$

(**246**) **Theorem:** Objects that Encode Properties Are Abstract. We can easily derive from the contraposition of axiom (52) that if x encodes a property, then x is abstract:

$$\exists F xF \to A!x$$

The converse fails because there exists an abstract *null* object that encodes no properties. See theorem (264.1) below.

(**247**) **Theorems:** Every Property is Encoded by Some Object and Every Relation is Encoded By Some Objects.

(.1) $\forall H \exists x\, xH$

(.2) $\forall G \exists x_1 \ldots \exists x_n (x_1 \ldots x_n G)$ $(n \geq 2)$

Moreover, given any n-ary relation term Π $(n \geq 1)$, it follows that Π is encoded by some objects just in case there is a relation that is identical to Π:

(.3) $\exists x\, x\Pi \equiv \exists H(H = \Pi)$, where Π is any unary relation term in which x and H don't occur free

(.4) $\exists x_1 \ldots \exists x_n (x_1 \ldots x_n \Pi) \equiv \exists G(G = \Pi)$, where Π is any n-ary relation term $(n \geq 2)$ in which $x_1, \ldots, x_n$ and G don't occur free

(**248**) **Remark:** Justifying the Definition of Property Existence. Note that the proof of (247.1) would have been much more involved if we had not defined $H{\downarrow}$ in (20.2) as $\exists x\, xH$. Without this definition, the proof of (247.1) requires an appeal to the Comprehension Principle for Abstract Objects (53):

By GEN, it suffices to prove $\exists x\, xH$. Note that the following is an instance of the Comprehension Principle for Abstract Objects (53):

$$\exists x(A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Hz)))$$

Suppose a is such an object, so that we know:

$$A!a \,\&\, \forall F(aF \equiv \forall z(Fz \equiv Hz))$$

Since $H\downarrow$ (39.2) and H is substitutable for F in the matrix of the second conjunct, we know $aH \equiv \forall z(Hz \equiv Hz)$. But $\forall z(Hz \equiv Hz)$ is an easily established theorem. Hence, it follows that aH, and so by $\exists$I, $\exists x\, xH$. ⋈

Similarly, without definition (20.2), the proof of (247.2), previewed in footnote 109 in item (33), is more complicated:

Again, by GEN, it suffices to prove $\exists x_1 \ldots \exists x_n(x_1 \ldots x_n G)$. Note that the following is an instance of the Comprehension Principle for Abstract Objects (53):

$$\exists x(A!x \,\&\, \forall F(xF \equiv F = F))$$

Suppose a is such an object, so that we know:

$$A!a \,\&\, \forall F(aF \equiv F = F))$$

Now it is straightforward to show that $\forall H\, aH$ (exercise). Note that by (39.2), all of the following properties exist: $[\lambda y\, Gya\ldots a]$, $[\lambda y\, Gaya\ldots a]$, …, and $[\lambda y\, Ga\ldots ay]$. So a encodes each of these properties. That is, we know:

$$a[\lambda y\, Gya\ldots a] \,\&\, a[\lambda y\, Gaya\ldots a] \,\&\, \ldots \,\&\, a[\lambda y\, Ga\ldots ay]$$

Hence, by axiom (50), $a\ldots aG$, and so $\exists x_1 \ldots \exists x_n(x_1 \ldots x_n G)$. ⋈

Since our system would have yielded these as theorems, it made sense to *define* the existence of properties and relations in terms of the fact that there are objects that encode them.

(**249**) **Exercise:** Practice with the Comprehension Principle for Abstract Objects. Say what is wrong with the following reasoning:

Consider any instance of the Comprehension Principle for Abstract Objects (53), say $\exists x(A!x \,\&\, \forall F(xF \equiv \varphi))$. Assume a is such an object, so that we know $A!a \,\&\, \forall F(aF \equiv \varphi)$. Hence, by RN, $\Box(A!a \,\&\, \forall F(aF \equiv \varphi))$. So by the usual sequence of $\exists$I and $\exists$E, it follows that $\exists x\Box(A!x \,\&\, \forall F(xF \equiv \varphi))$.

Now show that if the conclusion of this reasoning were correct, there would be a total modal collapse of the system, i.e., show that $\exists x\Box(A!x \,\&\, \forall F(xF \equiv \varphi))$ implies $\varphi \equiv \Box\varphi$, for an arbitrary formula φ.[172]

[172] Solution to the second part of the exercise: Suppose $\exists x\Box(A!x \,\&\, \forall F(xF \equiv \varphi))$, for an *arbitrary* formula φ. Then let a be such an object, so that we know $\Box(A!a \,\&\, \forall F(aF \equiv \varphi))$. Since a necessary conjunction implies that the conjuncts are necessary, it follows that $\Box A!a \,\&\, \Box\forall F(aF \equiv \varphi))$. From the second conjunct of this last result, it follows both that:

(**250**) **Theorems:** Strengthened Comprehension Principle for Abstract Objects. The Comprehension Principle for Abstract Objects (53) and the definition of identity (23.1) jointly imply that there is a *unique* abstract object that encodes just the properties such that φ:

$$\exists!x(A!x \,\&\, \forall F(xF \equiv \varphi)), \text{ provided } x \text{ doesn't occur free in } \varphi$$

The proof is simplified by appealing to (245.2).

(**251**) **Theorems:** Examples of Abstract Objects via Strengthened Comprehension. The Strengthened Comprehension Principle for Abstract Objects (250) asserts the unique existence of a number of interesting abstract objects. There exists a unique abstract object that encodes all and only the properties F such that: (.1) y exemplifies F; (.2) y and z exemplify F; (.3) y or z exemplify F; (.4) y necessarily exemplifies F; (.5) F is identical to property G; and (.6) F is necessarily implied by G:

(.1) $\exists!x(A!x \,\&\, \forall F(xF \equiv Fy))$

(.2) $\exists!x(A!x \,\&\, \forall F(xF \equiv Fy \,\&\, Fz))$

(.3) $\exists!x(A!x \,\&\, \forall F(xF \equiv Fy \vee Fz))$

(.4) $\exists!x(A!x \,\&\, \forall F(xF \equiv \Box Fy))$

(.5) $\exists!x(A!x \,\&\, \forall F(xF \equiv F = G))$

(.6) $\exists!x(A!x \,\&\, \forall F(xF \equiv \Box\forall y(Gy \rightarrow Fy)))$

Many of the above objects (and others) will figure prominently in the theorems which follow.

(**252**) **Theorems:** Descriptions Guaranteed to be Significant. We can now establish, for any condition φ in which x doesn't occur free, the existence of the abstract individual that encodes just the properties such that φ:

(ϑ) $\forall F\Box(aF \equiv \varphi)$ by the CBF schema (167.2)

(ξ) $\forall F(aF \equiv \varphi)$ by the T schema (45.2)

If we apply $\forall$E to both, we obtain, respectively:

(ϑ') $\Box(aF \equiv \varphi)$

(ξ') $aF \equiv \varphi$

From (ϑ') it follows by (158.6) that:

(ζ) $\Box aF \equiv \Box\varphi$

Now to see how this implies that truth and necessity are equivalent, we need only show $\varphi \rightarrow \Box\varphi$, since $\Box\varphi \rightarrow \varphi$ is an instance of the T schema. So assume φ. From this, it follows from (ξ') that aF. From this it follows that $\Box aF$, by axiom (51). So by (ζ), it follows that $\Box\varphi$.

$\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))\downarrow$, provided x doesn't occur free in φ

Although a non-modally strict proof is easily obtained from (250) and (144.1)★, we obtain a modally strict proof if we apply the Rule of Actualization to (176.2) (250) and then appeal to the right-to-left direction of theorem $(\imath x\varphi)\downarrow \equiv \mathcal{A}\exists!x\varphi$ (176.2).

Since the above theorem schema has a modally-strict proof, its necessitation follows by RN. If we speak intuitively in terms of primitive possible worlds, then note that the necessitation of this theorem does *not* say that, at every world $\boldsymbol{w}$, the x, which both exemplifies being abstract at $\boldsymbol{w}$ and encodes all and only the properties satisfying φ at $\boldsymbol{w}$, exists at $\boldsymbol{w}$. Rather, the necessitation says that at every possible world $\boldsymbol{w}$, the x, which both exemplifies being abstract at the distinguished actual world $\boldsymbol{w}_0$ and encodes exactly the properties satisfying φ at $\boldsymbol{w}_0$, exists at $\boldsymbol{w}$.

(253) **Metadefinitions:** Canonical Descriptions, Matrices, and Canonical Individuals. The previous theorem guarantees that descriptions of the form $\imath\nu(A!\nu \,\&\, \forall F(\nu F \equiv \varphi))$ are significant whenever ν is any individual variable that doesn't occur free in φ. We henceforth say:

> A definite description is *canonical* iff it has the form $\imath\nu(A!\nu \,\&\, \forall F(\nu F \equiv \varphi))$, for any formula φ in which the individual variable ν doesn't occur free.

We call the matrix $A!\nu \,\&\, \forall F(\nu F \equiv \varphi)$ of a canonical description a *canonical matrix*. By an abuse of language, we shall call the individuals denoted by such descriptions *canonical* individuals.

(254) ★**Theorems:** Canonical Individuals Encode Their Defining Properties. As an instance of (145.2)★, it is a theorem that if y is the abstract individual that encodes just the properties such that φ, then y is abstract and encodes just the properties such that φ:

$y = \imath x(A!x \,\&\, \forall F(xF \equiv \varphi)) \to (A!y \,\&\, \forall F(yF \equiv \varphi))$,
provided x doesn't occur free in φ

In general, we cannot give a modally strict proof of this claim. But as we shall see, for a certain group of formulas φ, there is a modally strict proof.

(255) **Theorems:** Canonical Individuals are Abstract. While the previous theorem implies that anything identical to a canonical individual exemplifies *being abstract*, this particular conditional can be derived by modally strict means:

$y = \imath x(A!x \,\&\, \forall F(xF \equiv \varphi)) \to A!y$, provided x doesn't occur free in φ

(256) ★**Theorems:** The Abstraction Principle for Canonical Individuals. It is a straightforwardly provable fact about canonical descriptions that the abstract object, which encodes exactly the properties such that φ, encodes F if and only if φ:

(.1) $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))F \equiv \varphi$, provided x doesn't occur free in φ

It is sometimes easier to understand (.1) if we derive it in a slightly different form. Note that the expression φ^G_F, i.e., the result of substituting G for every free occurrence of F in φ, asserts that G is such that φ. The Abstraction Principle for canonical individuals then implies that the abstract object, which encodes exactly the properties such that φ, encodes a property G if and only if G is such that φ:

(.2) $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G \equiv \varphi^G_F$,
provided x doesn't occur free in φ and G is substitutable for F in φ

To see an example, let b be an arbitrary object. Then we have the following instance of the above theorem: $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))G \equiv Gb$. This asserts: the abstract object that encodes exactly the properties that b exemplifies encodes G iff b exemplifies G. The reason we call this 'abstraction' should now be clear: in the right-to-left direction, we've abstracted out, from the simple predication Gb, an encoding claim about a particular abstract object. And (.1) is even more general: from any formula φ with no free xs, we may abstract out an encoding claim about a canonical individual.

We may describe this example intuitively as follows. When φ is formula Gb, then φ describes a unique logical pattern of predications about b, namely, the logical pattern consisting of those properties G that are such that Gb. The Strengthened Comprehension Principle (250) *objectifies* this pattern and asserts its unique existence. Moreover, (252) guarantees that the canonical description of the objectified pattern is significant. Finally, the Abstraction Principle then yields an important truth about this objectified pattern, namely, that it encodes G if and only G matches the pattern. This way of looking at the Abstraction Principle applies to arbitrarily complex formulas φ in which x doesn't occur free, since these express a condition on properties and thereby define a logical pattern of properties.

(**257**) **Remark:** The Abstraction Principle and Necessitation. Inspection shows that the proof of the Abstraction Principle depends upon Hintikka's schema (142)★, which in turn depends on the fundamental theorem (141)★ governing descriptions. So the Abstraction Principle is not a modally strict theorem and is not subject to the Rule of Necessitation. It would serve well to get a broad perspective on these important facts. In the discussion that follows, we provide such a perspective in the material mode and leave the explanation in the formal mode to a footnote.

To understand more fully why RN can't be applied to instances of Abstraction, suppose we want to extend our theory with some contingent facts, by asserting those facts as new axioms. For example, we might assert, as part of the body of truths we'd like to systematize with our theory, that the individual

b exemplifies being a philosopher but might not have been a philosopher. So, suppose we've extended our theory with the following two axioms, where P is the property *being a philosopher*:

> $\star$**Fact**: Pb
>
> **Modal Fact**: $\Diamond\neg Pb$

Thus, the claim Pb is contingently true, by definition (213.1). Hence Pb is a modally fragile axiom and that is why we've labeled Pb as a $\star$Fact. If these claims are taken as axioms, they become part of our deductive system and any reasoning that depends on $\star$Fact fails to be modally strict.

Now consider the following instance of the Abstraction Principle (256.2)$\star$:

(a) $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))P \equiv Pb$

It then follows from (a) and our $\star$Fact that:

(b) $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))P$

Since the description in (b) is canonical, it is significant (252) and so may be instantiated, along with P, in our axiom for the rigidity of encoding (51), which asserts $xF \rightarrow \Box xF$. By doing so, we may infer:

(c) $\Box\imath x(A!x \,\&\, \forall F(xF \equiv Fb))P$

Of course, the proof of (c) fails to be modally strict, since it was derived from *two* $\star$-claims: the above contingent $\star$Fact (Pb) and (a), which is an instance of (256.2)$\star$. Nevertheless, (c) is a theorem in the extended theory we're considering.

Now if we could apply RN to (a), we would obtain:

(d) $\Box(\imath x(A!x \,\&\, \forall F(xF \equiv Fb))P \equiv Pb)$

Then from (d), (c), and the relevant instance of (158.6), which asserts that $\Box(\varphi \equiv \psi) \rightarrow (\Box\varphi \equiv \Box\psi)$, it would follow that:

(e) $\Box Pb$

But this would contradict our Modal Fact $\Diamond\neg Pb$, which is equivalent to $\neg\Box Pb$. It should therefore be clear why RN isn't applicable to the Abstraction Principle — if it were so applicable, our system would become inconsistent when extended with natural but contingent (and thus modally fragile) axioms.[173]

[173] In the formal mode, the reason why we can't generally apply RN to instances of Abstraction is that there are interpretations in which the necessitation of the Abstraction Principle fails to be true, thus demonstrating its invalidity. To see this, let us again help ourselves for the moment to the semantically primitive notion of a possible world. The semantic counterpart of the above $\star$Fact

It is important to note that in the foregoing discussion, we have a described a scenario of the kind mentioned in (71), where we outlined conditions under which the converse of RN fails to hold. The converse of RN asserts that if $\vdash \Box\varphi$, then $\vdash_{\Box} \varphi$. But if we let φ be sentence (b) in the above scenario, then we indeed have $\vdash \Box\varphi$ and not $\vdash_{\Box} \varphi$. For we saw that since (b) is an encoding formula and a theorem, its necessitation (c) is also a theorem. But there isn't a modally strict proof of (b), since any proof would depend on our ⋆Fact Pb as well as some non-modally strict fact about descriptions, such as the right-to-left direction of the Abstraction Principle (256.2)⋆.[174]

(**258**) **Theorems:** Actualized Abstraction. By strategically placing an actuality operator in the Abstraction Principle, we obtain a principle that has a modally strict proof, in two forms, namely, (.1) the abstract object, which encodes just the properties such that φ, encodes a property F if and only if it is actually the case that φ, and (.2) the abstract object, which encodes just the properties such that φ, encodes a property G if and only if it is actually the case that G is such that φ:

---

and Modal Fact is any interpretation of our language in which Pb is true at the actual world $\boldsymbol{w}_0$ but false at some other possible world, say, $\boldsymbol{w}_1$. In such an interpretation, our instance (a) of the Abstraction Principle fails to be necessarily true. (a) fails to be necessarily true in the left-to-right direction by the following argument. Since Pb is true at $\boldsymbol{w}_0$, $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))$ encodes P at $\boldsymbol{w}_0$. Since the properties an object encodes are necessarily encoded (179.2), $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))$ encodes P at $\boldsymbol{w}_1$. But, by hypothesis, Pb is false at $\boldsymbol{w}_1$. Hence, $\boldsymbol{w}_1$ is a world that is the witness to the truth of the following possibility claim:

$$\Diamond(\imath x(A!x \,\&\, \forall F(xF \equiv Fb))P \,\&\, \neg Pb)$$

Since the left-to-right condition of (a) isn't necessary in this interpretation, the necessitation of this condition fails to be valid.

Similarly, if we consider the negation of P, namely $\overline{P}$, with respect to the interpretation described above, then the following consequence of the Abstraction Principle fails to be necessarily true in the right-to-left direction:

(f) $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))\overline{P} \equiv \overline{P}b$

To see that it is possible for the right condition of (f) to be true while the left condition false, we may reason as follows. Since Pb is true at $\boldsymbol{w}_0$ in the interpretation we've described, $\overline{P}b$ is false at $\boldsymbol{w}_0$ and so $\imath x(A!x \& \forall F(xF \equiv Fb))$ fails to encode $\overline{P}$ at $\boldsymbol{w}_0$. Since the properties an abstract object fails to encode are properties it necessarily fails to encode (179.7), it follows that $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))$ fails to encode $\overline{P}$ at $\boldsymbol{w}_1$. But, by hypothesis, Pb is false at $\boldsymbol{w}_1$ and so $\overline{P}b$ is true at $\boldsymbol{w}_1$. Hence, $\boldsymbol{w}_1$ is a witness to the truth of the following possibility claim:

$$\Diamond(\overline{P}b \,\&\, \neg\imath x(A!x \,\&\, \forall F(xF \equiv Fb))\overline{P})$$

Since this shows that the right-to-left direction of (f) isn't necessarily true in this interpretation, the necessitation of this direction fails to be valid.

[174]Since (b) is not an axiom, any proof of (b) must infer it by Modus Ponens from two previous lines in the proof. One of those lines has to be our ⋆Fact Pb and the other has to be the right-to-left direction of the Abstraction Principle (256.2)⋆. In other words, any proof of (b) must include the lines Pb (⋆Fact) and $Pb \to \imath x(A!x \,\&\, \forall F(xF \equiv Fb))P$ (256.2)⋆. But such lines aren't necessary truths. See the second part of footnote 173 for the reasoning that shows why the right-to-left direction of the Abstraction Principle isn't necessary.

(.1) $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))F \equiv \mathcal{A}\varphi$, provided x doesn't occur free in φ

(.2) $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G \equiv \mathcal{A}\varphi^G_F$,
provided x doesn't occur free in φ and G is substitutable for F in φ

The reader is encouraged to try proving these theorems without the benefit of the proof in the Appendix.

(**259**) **Theorems:** Properties That Are Necessarily Such That φ. There are modally strict proofs of the claims: (.1) if G is necessarily such that φ, then $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$ encodes G, and (.2) if G is necessarily such that φ, then necessarily, $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$ encodes G if and only if G is such that φ:

(.1) $\Box\varphi^G_F \rightarrow \imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G$,
provided x doesn't occur free in φ and G is substitutable for F in φ

(.2) $\Box\varphi^G_F \rightarrow \Box(\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G \equiv \varphi^G_F)$,
provided x doesn't occur free in φ and G is substitutable for F in φ

The proof of (.1) in the Appendix uses Actualized Abstraction (258.2), and the proof of (.2) utilizes (.1). Note that (.2) describes sufficient conditions under which instances of the Abstraction Principle, as formulated in (256.2)★, become necessary truths, namely, when G is a property that is necessarily such that φ.

Exercise: Give modally strict proofs of the simpler forms of these theorems, i.e.,

- $\Box\varphi \rightarrow \imath x(A!x \,\&\, \forall F(xF \equiv \varphi))F$, provided x doesn't occur free in φ
- $\Box\varphi \rightarrow \Box(\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))F \equiv \varphi)$, provided x doesn't occur free in φ

(**260**) **Metadefinitions:** Rigid Conditions and Strict Canonicity. We now introduce terms into our metalanguage to single out the special conditions under which canonical descriptions can be considered *strictly* canonical. Let φ be any formula in which the variable α may occur free. Then we say:

(.1) φ is a *rigid condition on* α if and only if $\vdash_\Box \forall\alpha(\varphi \rightarrow \Box\varphi)$.

Thus, if $\vdash_\Box \forall x(\varphi \rightarrow \Box\varphi)$, we say that φ is a rigid condition on objects, and if $\vdash_\Box \forall F(\varphi \rightarrow \Box\varphi)$, we say that φ is a rigid condition on properties. Thus, we may say:

(.2) The canonical description $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$ is *strictly canonical* just in case φ is a rigid condition on properties.

Henceforth, we sometimes abuse language and speak of the strictly canonical individuals denoted by strictly canonical descriptions.

(**261**) **Theorems:** Facts About Strict Canonicity. When φ is a rigid condition on properties, it is a modally strict fact that (.1) if x is an abstract individual that encodes exactly the properties such that φ, then necessarily x is an abstract individual that encodes exactly the properties such that φ:

(.1) $(A!x \& \forall F(xF \equiv \varphi)) \rightarrow \Box(A!x \& \forall F(xF \equiv \varphi))$, provided φ is a rigid condition on properties in which x doesn't occur free.

Moreover, when φ is a rigid condition on properties, it is a modally strict fact that (.2) anything identical to a strictly canonical individual both exemplifies *being abstract* and encodes all and only the properties such that φ:

(.2) $y = \imath x(A!x \;\&\; \forall F(xF \equiv \varphi)) \rightarrow (A!y \;\&\; \forall F(yF \equiv \varphi))$, provided φ is a rigid condition on properties in which x doesn't occur free.

Thus, theorem (254)★ has a modally strict special case, namely, when the description $\imath x(A!x \;\&\; \forall F(xF \equiv \varphi))$ is strictly canonical.

Finally, we note that the Abstraction Principle formulated in (256.1)★ becomes a modally strict theorem with respect to strictly canonical objects:

(.3) $\imath x(A!x \;\&\; \forall F(xF \equiv \varphi))F \equiv \varphi$, provided φ is a rigid condition on properties in which x doesn't occur free.

As it turns out, however, (.2) will be more useful than (.3). Since we'll introduce many new terms using a definition-by-identity in which the definiens is a strictly canonical description, (.2) becomes immediately applicable to the identity statements such definitions make available as theorems.

(**262**) **Remark:** Some Examples of Strictly Canonical Individuals. In item (251), we presented a series of instances of the Strengthened Comprehension Principle for Abstract Objects. The last three are examples of strictly canonical individuals. Consider the following canonical descriptions based on those last three examples:

(a) $\imath x(A!x \;\&\; \forall F(xF \equiv \Box Fy))$

(b) $\imath x(A!x \;\&\; \forall F(xF \equiv F = G))$

(c) $\imath x(A!x \;\&\; \forall F(xF \equiv \Box\forall y(Gy \rightarrow Fy)))$

These are all instances of $\imath x(A!x \& \forall F(xF \equiv \varphi))$ and, hence, significant, by (252). Moreover, in each case, the formula φ in question is a rigid condition on properties, i.e., $\vdash_\Box \forall F(\varphi \rightarrow \Box\varphi)$. We can see this as follows:

- In example (a), φ is $\Box Fy$. But as an instance of the 4 schema (165.5), we know $\Box Fy \to \Box\Box Fy$, i.e., $\varphi \to \Box\varphi$. Hence, by applying GEN, we have $\forall F(\varphi \to \Box\varphi)$.

- In example (b), φ is $F = G$, and so by (125.1), we know that $\varphi \to \Box\varphi$. So, again, by GEN we have $\forall F(\varphi \to \Box\varphi)$.

- In example (c), φ is $\Box\forall y(Gy \to Fy)$. Hence we can reason as we did in (a), by way of the 4 schema, to establish $\forall F(\varphi \to \Box\varphi)$.

Thus, (a) – (c) are examples of strictly canonical individuals.

(**263**) **Definitions:** Null and Universal Objects. We say: (.1) x is a *null* object just in case x is an abstract object that encodes no properties; and (.2) x is a *universal* object just in case x is an abstract object that encodes every property:

(.1) $Null(x) \equiv_{df} A!x \,\&\, \neg\exists FxF$

(.2) $Universal(x) \equiv_{df} A!x \,\&\, \forall FxF$

We are taking advantage here of our conventions in (17.2) by using object language variables as the free variables in the definiens and definiendum, instead of metavariables. Moreover, since for any individual term κ, the claim $A!\kappa$ implies $\kappa{\downarrow}$, no existence clauses have to be added to the definiens for (.1) and (.2), for the reasons noted in Remark (36).

(**264**) **Theorems:** Existence and Uniqueness of Null and Universal Objects. It is now easily established that (.1) there is a unique null object, and (.2) there is a unique universal object:

(.1) $\exists!xNull(x)$

(.2) $\exists!xUniversal(x)$

Consequently, it follows, by a modally strict proof that (.3) the null object exists, and (.4) the universal object exists:

(.3) $\imath xNull(x){\downarrow}$

(.4) $\imath xUniversal(x){\downarrow}$

(**265**) **Definitions:** Notation for the Null and Universal Objects. We now introduce new constants to designate the null object and the universal object:

(.1) $\boldsymbol{a}_{\varnothing} =_{df} \imath xNull(x)$

(.2) $\boldsymbol{a}_{\boldsymbol{V}} =_{df} \imath xUniversal(x)$

Since $\iota x Null(x)$ and $\iota x Universal(x)$ are provably significant, the Rule of Identity by Definition (120.1) tells us that definitions become theorems asserting identities.

(**266**) **Theorems:** Facts Implied by the Foregoing. We begin with some facts about the conditions $Null(x)$ and $Universal(x)$: (.1) if x is a null object, then necessarily x is a null object, and (.2) if x is a universal object, then necessarily x is a universal object:

(.1) $Null(x) \rightarrow \Box Null(x)$

(.2) $Universal(x) \rightarrow \Box Universal(x)$

We may also prove, by modally strict means, that (.3) the null object is a null object; (.4) the universal object is a universal object; (.5) the null object is not identical to the universal object; (.6) the null object is identical to the (canonical) abstract object that encodes all and only non-self-identical properties; and (.7) the universal object is identical to the (canonical) abstract object that encodes all and only self-identical properties:

(.3) $Null(\boldsymbol{a}_\varnothing)$

(.4) $Universal(\boldsymbol{a}_V)$

(.5) $\boldsymbol{a}_\varnothing \neq \boldsymbol{a}_V$

(.6) $\boldsymbol{a}_\varnothing = \iota x(A!x \,\&\, \forall F(xF \equiv F \neq F))$

(.7) $\boldsymbol{a}_V = \iota x(A!x \,\&\, \forall F(xF \equiv F = F))$

Though these theorems sound trivial, their proof by modally strict means is not.

(**267**) **Remark:** A Rejected Alternative. Now that we've seen how to obtain a modally strict proof of (266.3) and (266.6), one might wonder whether it would have been simpler to define $\boldsymbol{a}_\varnothing$ directly as:

(ϑ) $\boldsymbol{a}_\varnothing =_{df} \iota x(A!x \,\&\, \forall F(xF \equiv F \neq F))$

This alternative definition (ϑ) is certainly legitimate, but we have a good reason for not deploying it. Since (ϑ) defines $\boldsymbol{a}_\varnothing$ as the abstract object that encodes all and only non-self-identical properties, it does not explicitly introduce $\boldsymbol{a}_\varnothing$ as *the null object*. Of course, one can use (ϑ) to derive that $\boldsymbol{a}_\varnothing$ is a *unique* null object, i.e., that both $Null(\boldsymbol{a}_\varnothing)$ and $\forall y(Null(y) \rightarrow y = \boldsymbol{a}_\varnothing)$.[175] One could then start

[175]Let φ be the formula $F \neq F$ and let χ be the formula $A!x \,\&\, \forall F(xF \equiv F \neq F)$. Now by applying GEN to an appropriate instance of (170.2), we have a modally strict proof of $\forall F(\varphi \rightarrow \Box\varphi)$. Since φ is therefore a rigid condition on properties (260.1), (ϑ) defines $\boldsymbol{a}_\varnothing$ as a strictly canonical object. So the identity $\boldsymbol{a}_\varnothing = \iota x\chi$ licensed by (ϑ) implies $\chi^{\boldsymbol{a}_\varnothing}_x$, i.e., $A!\boldsymbol{a}_\varnothing \& \forall F(\boldsymbol{a}_\varnothing F \equiv F \neq F)$, by (261.2). But from this we can derive $Null(\boldsymbol{a}_\varnothing)$ since all that remains to be shown is that the second conjunct implies $\neg\exists F \boldsymbol{a}_\varnothing F$ (exercise). Moreover, it is then easy to establish that $\forall y(Null(y) \rightarrow y = \boldsymbol{a}_\varnothing)$ (exercise).

referencing it as 'the null object'. But this fails to take advantage of the fact that our system allows us to properly define the condition $Null(x)$ as $A!x \,\&\, \neg\exists F xF$, prove $\exists! x Null(x)$, show $\imath x Null(x)\downarrow$, and then use the well-defined description $\imath x Null(x)$ to introduce the notation $\boldsymbol{a}_\varnothing$. If we are going to use the expression $\boldsymbol{a}_\varnothing$ to name the null object, the correct way to do so is to define $\boldsymbol{a}_\varnothing$ as $\imath x Null(x)$ after having shown that the latter is significant. But once we do this, it takes a little work to show that $\boldsymbol{a}_\varnothing$ can be instantiated into its own defining matrix by a modally strict proof.

A similar observation may be made about whether it might have been simpler to define $\boldsymbol{a}_{\boldsymbol{V}}$ directly as $\imath x(A!x \,\&\, \forall F(xF \equiv F = F))$.

(**268**) **Theorems:** Facts About the Granularity of Relations. The following facts govern the granularity of relational properties having abstract constituents: (.1) for any binary relation F, there are distinct abstract objects x and y for which *bearing F to x* is identical to *bearing F to y*, and (.2) for any binary relation F, there are distinct abstract objects x and y for which *being a z such that x bears F to z* is identical to *being a z such that y bears F to z*:

(.1) $\forall F \exists x \exists y(A!x \,\&\, A!y \,\&\, x \neq y \,\&\, [\lambda z\, Fzx] = [\lambda z\, Fzy])$

(.2) $\forall F \exists x \exists y(A!x \,\&\, A!y \,\&\, x \neq y \,\&\, [\lambda z\, Fxz] = [\lambda z\, Fyz])$

These theorems are to be expected if we think semantically for the moment and reconsider the Aczel models discussed in Chapter 3. Recall that for the purpose of modeling the theory, abstract objects may be represented as sets of properties (though for the reasons pointed out earlier, we shouldn't confuse abstract objects with the sets that represent them). Now consider any relation R. Cantor's Theorem now tells us there can't be a distinct property *bearing F to s* for each distinct set s of properties, for then we would have a one-to-one mapping from the power set of the set of properties into a subset of the set of properties. This model-theoretic fact is captured by the above theorem: there can't be a distinct property *bearing F to x* for each distinct abstract object x. Thus, Cantor's Theorem isn't violated: there are so many abstract objects that for some distinct abstract objects x and y, the property *bearing F to x* collapses to the property *bearing F to y*. This is just what one expects given that two powerful principles, comprehension for abstract objects (53) and comprehension for properties (191.2), are true simultaneously.

It also follows that, for any property F, there are distinct abstract objects x, y such that *that-Fx* is identical to *that-Fy*:

(.3) $\forall F \exists x \exists y(A!x \,\&\, A!y \,\&\, x \neq y \,\&\, [\lambda\, Fx] = [\lambda\, Fy])$

This is as expected, given (.1) and (.2).

(**269**) **Theorem:** Abstract Objects Are Not Strictly Leibnizian. A previous theorem, (268.1), has a rather interesting consequence, namely, that there exist distinct abstract objects that exemplify exactly the same properties:[176]

$$\exists x \exists y (A!x \,\&\, A!y \,\&\, x \neq y \,\&\, \forall F(Fx \equiv Fy))$$

In other words, there are distinct abstract objects that are indiscernible with respect to the properties they exemplify.[177] The proof of this claim begins by instantiating theorem (268.1) to the relation $[\lambda xy\, \forall F(Fx \equiv Fy)]$.

Thus classical Leibnizian indiscernibility doesn't imply the identity of abstract objects. Nevertheless, in light of theorem (245.2), a related form of the identity of indiscernibles applies to abstract objects: such objects are identical whenever they are indiscernible with respect to the properties they encode.

9.11.3 Discernible Objects

Theorem (269) opens Pandora's Box[178] by establishing that there are objects that fail Leibniz's principle of the identity of indiscernibles. But one can easily carve out a domain of *discernible* objects that obey Leibniz's principle. The definition is given in (273.2) below, and though a discernible object can be ordinary or abstract, we won't, as yet, be able to prove that there are discernible abstract objects. This will be established later, in Chapter 14, once we extend object theory with a single new axiom.

Before we introduce the definition of discernibility, we first prove a theorem that plays an important role in establishing the existence of properties and relations. Intuitively, the theorem tells us that φ can serve as matrix for a significant λ-expression just in case φ can't distinguish between indiscernible objects, i.e., just in case φ can't distinguish between objects that exemplify the same properties. This theorem was contributed by Daniel Kirchner and it shows an interesting duality: on the one hand, discernibles are distinguished by the pattern of their property-exemplifications, and on the other hand, a property exists whenever it can't distinguish indiscernibles.

(**270**) **Remark:** In Preparation for Kirchner's Theorem. By deploying tools developed by, and in collaboration with, Christoph Benzmüller, Daniel Kirchner

[176] I am indebted to Peter Aczel for pointing out (in email correspondence, 11 November 1996) that this theorem results once impredicatively-defined relations are allowed into the system.

[177] If one recalls the structure of the Aczel models, this result is to be expected. When properties are modeled as sets of urelements, and abstract objects are modeled as sets of properties, then the claim that an abstract object x exemplifies a property F is true just in case the special urelement that serves as the proxy of x is an element of F. Consequently, since distinct abstract objects must sometimes have the same proxy, some distinct abstract objects exemplify the same properties.

[178] Thanks to Uri Nodelman for noticing the appropriateness of this metaphor.

has been implementing object theory in Isabelle/HOL.[179] This implementation consists of two parts: (a) a *shallow semantic embedding* of object theory in Isabelle/HOL, and (b) a reconstruction of PLM's deductive system. (a) The shallow semantic embedding uses the infrastructure of Isabelle/HOL (which includes set theory in the framework of functional type theory) to (i) construct enhanced Aczel models of object theory, (ii) reconstruct (i.e., define) the syntax of object theory in terms of this model (by preserving and using as much of the syntax of Isabelle/HOL as possible),[180] and (iii) prove that the axioms of object theory, thus reconstructed, are true in the model. (b) Using the shallow semantic embedding as a basis, Kirchner implements object theory computationally by rebuilding its deductive system, as developed here, on top of the reconstructed axioms. Thus, the automated/interactive reasoning system that results doesn't imply any *artifactual* theorems, i.e., the resulting proof system doesn't imply any set-theoretic consequences of the Aczel model that are not expressible in the language of object theory or any consequences that are expressible but not derivable in object theory.

The two key principles of Aczel's models of object theory are that (a) properties can be modeled as sets of urelements (and relations as sets of *n*-tuples of urelements), and (b) abstract objects can be modeled as sets of properties. However, the urelements include a subdomain of special *proxy* individuals that *stand in for* abstract objects when evaluating exemplification formulas. As noted in Chapter 3, distinct abstract objects must sometimes have the same proxy (given that there are more sets of properties than there are urelements), and so Aczel models demonstrate why there are distinct but indiscernible abstract objects. Only encoding formulas can discriminate among distinct but indiscernible abstract objects and these formulas can't always express new properties or relations on pain of contradiction, for otherwise, we could establish a one-to-one correlation from the set of abstract objects (i.e., the set of sets of properties) to a subset of the set of properties, in violation of Cantor's Theorem.

Therefore, in building the shallow semantic embedding of object theory in

[179] See Kirchner 2017 [2021], Kirchner 2022, and Kirchner's GitHub repository at https://github.com/ekpyron/TAO/tree/NewAOT; see also Kirchner, Benzmüller, & Zalta 2020.

[180] Automated reasoning in object theory, to derive consequences of premises, takes place directly with respect to the formulas so defined, which are subject to the logic implemented in Isabelle/HOL. This stands in contrast to a *deep semantic embedding*, in which (i) each element of the language of object theory would be represented in Isabelle/HOL in terms of an inductive data structure that realizes the BNF, and (ii) the semantics of object theory would be stated, and a formula would be evaluated, by recursively traversing the data structure and translating it into the language of Isabelle/HOL. In a deep semantic embedding, automated reasoning about object theory, to derive theorems and consequences of premises, would take place only after each formula is semantically evaluated in Isabelle/HOL's infrastructure, since those semantic evaluations are governed by the logic implemented in Isabelle/HOL.

Isabelle/HOL, Kirchner's construction of enhanced Aczel models is based on a principle governing the conditions under which a formula φ expresses a property or relation, i.e., may serve as a matrix for a significant λ-expression (or as a matrix for the comprehension principle for relations). To repeat, the idea is that φ is a matrix for a significant λ-expression just in case φ can't distinguish between indiscernible objects, i.e., φ can't distinguish between objects that exemplify the same properties. When the current formulations of axioms (39.2) and (49) were put into place, Kirchner found a proof of his principle in the theory itself.[181] We present this theorem next; the proof in the Appendix was developed on the basis of Kirchner's proof sketch.

(**271**) **Theorems:** Kirchner's Theorem. In the unary case, Kirchner's Theorem asserts that a λ-expression of the form $[\lambda x\, \varphi]$ signifies a property just in case necessarily, for any indiscernible objects x and y (i.e., for any objects that exemplify the same properties), φ is equivalent to φ^{y}_{x}. The theorem can be stated formally as follows:

(.1) $[\lambda x\, \varphi]\downarrow \equiv \Box\forall x\forall y(\forall F(Fx \equiv Fy) \to (\varphi \equiv \varphi^{y}_{x}))$,
provided y doesn't occur free in φ.

Thus, the right-to-left direction tells us that as long as it is a necessary fact that φ can't distinguish among indiscernible objects, then $[\lambda x\, \varphi]$ is significant, even if it isn't a core λ-expression (i.e., even if the λ binds an occurrence of x in encoding position in φ).

Kirchner's Theorem generalizes to:

(.2) $[\lambda x_1 \ldots x_n\, \varphi]\downarrow \equiv$
$\Box\forall x_1 \ldots \forall x_n \forall y_1 \ldots \forall y_n(\forall F(Fx_1 \ldots x_n \equiv Fy_1 \ldots y_n) \to (\varphi \equiv \varphi^{y_1,\ldots,y_n}_{x_1,\ldots,x_n}))$,
provided $y_1,\ldots,y_n$ don't occur free in φ $(n \geq 1)$

In other words, $[\lambda x_1 \ldots x_n\, \varphi]$ exists just in case necessarily, φ and $\varphi^{y_1,\ldots,y_n}_{x_1,\ldots,x_n}$ are equivalent with respect to any objects $x_1,\ldots,x_n$ and $y_1,\ldots,y_n$ that exemplify the same relations.

To see (.1) in action, consider that its contrapositive implies (.3) there are properties G such that $[\lambda x\, xG]$ fails to signify a property, i.e.,

(.3) $\exists G\neg([\lambda x\, xG]\downarrow)$

It is not hard to find a witness, i.e., a property G for which the formula xG $(= \varphi)$ fails the right-hand condition of (.1), thereby allowing us to infer $\neg([\lambda x\, xG]\downarrow)$. To see this, note that by (269), there are objects, say a and b, such that $A!a$, $A!b$, $a \neq b$, and $\forall F(Fa \equiv Fb)$. Since a and b are distinct abstract objects, one encodes a property that the other doesn't encode. Without loss of generality,

[181] Personal email communication, 3 March 2020.

let G be a property such that aG and $\neg bG$. Then we know $\neg(aG \equiv bG)$. So we have $\forall F(Fa \equiv Fb) \rightarrow \neg(aG \equiv bG)$. Hence, $\exists x \exists y(\forall F(Fx \equiv Fy) \,\&\, \neg(xG \equiv yG))$. So $\Diamond \exists x \exists y(\forall F(Fx \equiv Fy) \,\&\, \neg(xG \equiv yG))$. Hence, the right side of (.1) is false, and so it follows that $\neg[\lambda x\, xG]\downarrow$. This holds for any G that distinguishes indiscernible abstract objects by virtue of what they encode.

(272) **Theorems:** Corollaries to Kirchner's Theorem. It follows from Kirchner's Theorem and the fact that indiscernibles are necessarily indiscernible (241) that if a λ-expression of the form $[\lambda x\, \varphi]$ signifies a property, then φ and φ^y_x are necessarily equivalent with respect to any indiscernible objects x and y:

(.1) $[\lambda x\, \varphi]\downarrow \rightarrow \forall x \forall y(\forall F(Fx \equiv Fy) \rightarrow \Box(\varphi \equiv \varphi^y_x))$,
provided y doesn't occur free in φ.

And the theorem generalizes to:

(.2) $[\lambda x_1 \ldots x_n\, \varphi]\downarrow \rightarrow$
$\forall x_1 \ldots \forall x_n \forall y_1 \ldots \forall y_n(\forall F(Fx_1 \ldots x_n \equiv Fy_1 \ldots y_n) \rightarrow \Box(\varphi \equiv \varphi^{y_1,\ldots,y_n}_{x_1,\ldots,x_n}))$,
provided $y_1,\ldots,y_n$ don't occur free in φ $(n \geq 1)$

In other words, if a λ-expression of the form $[\lambda x_1 \ldots x_n\, \varphi]$ signifies a relation, then φ and $\varphi^{y_1,\ldots,y_n}_{x_1,\ldots,x_n}$ are necessarily equivalent with respect to any objects $x_1,\ldots,x_n$ and $y_1,\ldots,y_n$ that exemplify the same relations.

As straightforward consequences of the Corollary to the Kirchner Theorem, it follows that (.3) whenever x and y are indiscernible objects, then *being a z such that x bears G to z* is identical to *being a z such that y bears G to z*, and (.4) whenever x and y are indiscernible objects, then *being a z such that z bears G to x* is identical to *being a z such that z bears G to y*:[182]

(.3) $\forall F(Fx \equiv Fy) \rightarrow [\lambda z\, Gxz] = [\lambda z\, Gyz]$

(.4) $\forall F(Fx \equiv Fy) \rightarrow [\lambda z\, Gzx] = [\lambda z\, Gzy]$

Note that this generalizes to any significant term $[\lambda z\, \varphi]$ that meets a well-defined proviso:

(.5) $[\lambda z \varphi]\downarrow \rightarrow (\forall F(Fx \equiv Fy) \rightarrow [\lambda z \varphi] = [\lambda z \varphi^y_x])$, provided that none of the free occurrences of x in φ are in encoding position.

Exercise: Show that $\forall F(Fx \equiv Fy) \rightarrow \forall G(Gx = Gy)$, i.e., that if x and y are indiscernible, then for any property G, (the proposition) that x exemplifies G is identical to (the proposition) that y exemplifies G. (A proof is given in the Appendix, after the proof of (.5).)

[182]Thanks to Daniel Kirchner for noting these consequences. They were artifactual theorems of the his Isabelle/HOL model of object theory, but became *bona fide* theorems of object theory once the final form of axiom (39.2) was reached.

(273) **Definitions and Theorems:** Discernible Objects. We have been informally using the notion *indiscernibles* to describe objects x and y such that $\forall F(Fx \equiv Fy)$. We could, of course, easily say when objects x and y are discernible, namely, when they fail to be indiscernible, i.e., when they exemplify the relation $[\lambda xy\, \neg\forall F(Fx \equiv Fy)]$. But this defines discernibility as a binary *relation*. In what follows, we want to define a *property* of discernibility. Intuitively, the idea is that an object x is discernible just in case, necessarily, for any distinct object y, some property distinguishes x and y.

To implement this idea, we first establish that (.1) [the property] *being an x such that necessarily, for every other object z, some exemplified property distinguishes x and z* exists:

(.1) $[\lambda x\, \Box\forall z(z \neq x \rightarrow \exists F\neg(Fz \equiv Fx))]\downarrow$

The proof of (.1) relies on Kirchner's Theorem. Though we might express (.1) more intuitively as *necessarily being distinguished from every other object by some exemplified property*, this is *not* to say that necessarily, there is some property distinguishes x from every other object (the claim $\forall z(z \neq x \rightarrow \exists F\varphi)$ doesn't imply $\exists F\forall z(z \neq x \rightarrow \varphi)$, for arbitrary φ). We may therefore define *being discernible* ('$D!$') in terms of property guaranteed to exist by (.1):

(.2) $D! =_{df} [\lambda x\, \Box\forall z(z \neq x \rightarrow \exists F\neg(Fz \equiv Fx))]$

Clearly, '$D!$' is well-defined; it follows from (.1) and (.2) that $D!\downarrow$. We include the modal operator $\Box$ in the definition because discernibility is a modal notion, in the first instance. It may help to see how (.2) captures this idea by noting that it implies: $D!x$ if and only if is *not possible* for there to be an object y distinct from x that exemplifies exactly the same properties as x.[183] However, since we know, by theorem (241) that $\forall F(Fx \equiv Fy) \rightarrow \Box\forall F(Fx \equiv Fy)$, it follows that an object is discernible if and only if it is, *in fact*, distinguishable from every other object:

(.3) $D!x \equiv \forall z(z \neq x \rightarrow \exists F\neg(Fz \equiv Fx))$

So we need not show that x is *necessarily* distinguishable from from every other object to show that it is a discernible object; the mere fact that it is distinguishable from every other object suffices.

In what follows then, we shall use *discernible* to refer to the property of objects that we just defined, but we'll continue to use *indiscernibles* to refer to the binary relation that holds between objects x and y whenever they exemplify the same properties. So the reader is forewarned that indiscernibility is not the

[183] For the ($\rightarrow$) direction, assume $D!x$. Then by (.2) and λ-Conversion, $\Box\forall z(z \neq x \rightarrow \exists F\neg(Fz \equiv Fx))$. This is equivalent to $\neg\Diamond\exists z\neg(z \neq x \rightarrow \exists F\neg(Fz \equiv Fx))$. In turn, this is equivalent to: $\neg\Diamond\exists z(z \neq x\ \&\ \forall F(Fz \equiv Fx))$. For the ($\leftarrow$) direction, reverse the reasoning.

negation of discernibility, though one could easily formulate a *second* notion of indiscernibility, call it *non-discernibility*, as a property that is the negation of $D!$. But we shouldn't need this notion in what follows.

Clearly, it now follows that (.4) ordinary objects are discernible, (.5) there are discernible objects; and (.6) some abstract objects are not discernible:

(.4) $O!x \to D!x$

(.5) $\exists x D!x$

(.6) $\exists x(A!x \,\&\, \neg D!x)$

It is also straightforward to show that (.7) indiscernible discernible objects are identical:

(.7) $(D!x \vee D!y) \to (\forall F(Fx \equiv Fy) \to x = y)$

Furthermore, without much effort, it can be shown that an object is necessarily discernible if discernible:

(.8) $D!x \to \Box D!x$

Hence, by RN, $\Box(D!x \to \Box D!x)$, and so $D!x$ is a modally collapsed formula. Thus, all the usual consequences of modal collapse are derivable, namely that (.9) x is discernible iff necessarily discernible; (.10) x is possibly discernible iff discernible; (.11) x is possibly discernible iff necessarily discernible; and (.12) x is discernible iff actually discernible:

(.9) $D!x \equiv \Box D!x$

(.10) $\Diamond D!x \equiv D!x$

(.11) $\Diamond D!x \equiv \Box D!x$

(.12) $D!x \equiv \mathcal{A} D!x$

We next show that [the property] *being a discernible object such that* φ exists, for any formula φ:

(.13) $[\lambda x\, D!x \,\&\, \varphi]\!\downarrow$, for any formula φ

However, to establish the more general claim, namely that:

$$[\lambda x_1 \ldots x_n\, D!x_1 \,\&\, \ldots \,\&\, D!x_n \,\&\, \varphi]\!\downarrow$$

it helps to have the following Lemma, which tells us that if objects $x_1,\ldots,x_n$ and $z_1,\ldots,z_n$ stand in the same relations, then x_1 and z_1 are indiscernible, and ... , and x_n and z_n are indiscernible:

(.14) $\forall F(Fx_1 \dots x_n \equiv Fz_1 \dots z_n) \rightarrow \forall G(Gx_1 \equiv Gz_1) \,\&\, \dots \,\&\, \forall G(Gx_n \equiv Gz_n) \quad (n \geq 1)$

With this lemma, we now establish that [the relation] *being objects* $x_1, \dots, x_n$ *that are discernible and such that* φ exists, for any φ:

(.15) $[\lambda x_1 \dots x_n\, D!x_1 \,\&\, \dots \,\&\, D!x_n \,\&\, \varphi]\!\downarrow$, for any formula φ $\quad (n \geq 1)$

Thus, as an instance of (.15), it is a theorem that (.16) *being discernible objects* x *and* y *that are identical* exists:

(.16) $[\lambda xy\, D!x \,\&\, D!y \,\&\, x = y]\!\downarrow$

Hence, we may define (.1) *identity for discernibles* ('$=_D$') as the binary relation *being identical discernibles* :

(.17) $=_D \;=_{df}\; [\lambda xy\, D!x \,\&\, D!y \,\&\, x = y]$

Clearly, then, $=_D$ exists. We henceforth use $=_D$ in infix notation. Hence, it follows immediately that (.18) x and y are identical$_D$ if and only if they are identical discernible objects; (.19) if x and y are identical$_D$ then they are identical; and (.20) x and y are identical$_D$ if and only if they are discernible objects that are necessarily indiscernible:

(.18) $x =_D y \equiv D!x \,\&\, D!y \,\&\, x = y$

(.19) $x =_D y \rightarrow x = y$

(.20) $x =_D y \equiv D!x \,\&\, D!y \,\&\, \Box\forall F(Fx \equiv Fy)$

We next prove that identity$_D$ is modally collapsed and all the usual consequences of that fact. Instead of first proving that $x =_D y \rightarrow \Box(x =_D y)$ is a modally strict theorem (which is sufficient, by RN, to establish that $x =_D y$ is modally collapsed), we first prove that (.21) x and y are identical$_D$ if and only if it is necessary that they are identical$_D$; (.22) it is possible that x and y are identical$_D$ if and only if they are identical$_D$; (.23) possibly x and y are identical$_D$ iff necessarily x and y are identical$_D$; and (.24) x and y are identical$_D$ iff it is actually the case that x and y are identical$_D$:

(.21) $x =_D y \equiv \Box(x =_D y)$

(.22) $\Diamond(x =_D y) \equiv x =_D y$

(.23) $\Diamond(x =_D y) \equiv \Box(x =_D y)$

(.24) $x =_D y \equiv \mathcal{A}(x =_D y)$

Now let us adopt the same conventions for expressing facts about the negation of $=_D$ as those used for expressing facts about the negation of $=_E$. We write $\neq_D$ instead of $\overline{=_D}$, and where κ_1 and κ_2 are any individual terms, we henceforth write $\kappa_1 \neq_D \kappa_2$ instead of $\neq_D \kappa_1 \kappa_2$. Hence it follows that:

(.25) $x \neq_D y \equiv \neg(x =_D y)$

The remarks about $\neq_E$ (236) now apply to $\neq_D$, for (.25) may also fail for arbitrary individual terms κ and κ': the schemas $\kappa \neq_D \kappa'$ and $\neg(\kappa =_D \kappa')$ are *not* equivalent. For if either $\neg\kappa\downarrow$ or $\neg\kappa'\downarrow$ is provable, then both $\kappa =_D \kappa'$ and $\kappa \neq_D \kappa'$ are provably false (they are both exemplification formulas with an empty primary term), in which case $\neg(\kappa =_E \kappa')$ becomes provably true.

Moreover, the modal collapse of $x \neq_D y$ and its consequences hold:

(.26) $x \neq_D y \equiv \Box x \neq_D y$

(.27) $\Diamond x \neq_D y \equiv x \neq_E y$

(.28) $\Diamond x \neq_E y \equiv \Box x \neq_E y$

(.29) $x \neq_E y \equiv \mathcal{A} x \neq_E y$

We next show that that $=_D$ is (.30) reflexive with respect to the discernible objects, (.31) symmetric, and (.32) transitive:

(.30) $D!x \rightarrow x =_D x$

(.31) $x =_D y \rightarrow y =_D x$

(.32) $(x =_D y \;\&\; y =_D z) \rightarrow x =_D z$

Finally, we conclude with some facts about $D!$, $=_D$, and $=$, namely that (.33) if x or y is discernible, then necessarily, x and y are identical just in case they are identical$_D$; (.34) if y is discernible, then its haecceity exists; and (.35) distinct discernibles have distinct haecceities:

(.33) $(D!x \vee D!y) \rightarrow \Box(x = y \equiv x =_D y)$

(.34) $D!y \rightarrow [\lambda x\, x = y]\downarrow$

(.35) $(D!x \;\&\; D!y) \rightarrow (x \neq y \equiv [\lambda z\, z = x] \neq [\lambda z\, z = y])$

(**274**) **Remark:** Why Discernible Objects are Philosophically Interesting. At present, we haven't asserted any axioms that imply that there are discernible abstract objects. In absence of any further axioms or applications (i.e., new primitive properties and axioms governing those properties), the theory is consistent with the claim that all abstract objects are indiscernible. (Indeed, this is the case in the smallest models of the theory.) So, for the moment, ordinary objects will serve well enough as paradigm examples of discernible objects. This, of course, need not remain true once the theory is extended or applied. In particular, in Chapter 14, we'll assert an axiom that implies the existence of discernible abstract objects.

The primary philosophical interest of discernible objects, however, is to put the lid back on Pandora's Box, which was opened by a theorem (269) that proved the existence of non-Leibnizian objects (i.e., distinct objects that are indiscernible). But much of philosophy assumes that objects are Leibnizian. In classical second-order logic without identity (i.e., with no encoding formulas), one typically defines object identity $x = y$ as $\forall F(Fx \equiv Fy)$; moreover, every condition on objects (i.e., every formula in which x occurs free) defines a property.[184] In the present system, every condition on *discernible* objects (i.e., every formula of the form $D!x \,\&\, \varphi$, whether or not x occurs free in φ) defines a property (273.13).

So a traditional assumption of classical logic, namely that individuals are governed by the identity of indiscernibles, is preserved with respect to discernible objects. Indeed, we may regard the Kirchner Theorem (271.1) and (271.2), and theorems (273.9) and (273.11), as establishing that relation comprehension is unrestricted on the domain of discernible objects and, thus, that relation and property comprehension in (191.1) and (191.2) extends classical second-order comprehension with new properties that behave exactly as expected on discernible objects.

9.12 Propositional Properties

(**275**) **Definition:** Propositional Properties. Let us say that a property F is *propositional* iff for some proposition p, F is *being such that p*:

$$Propositional(F) \equiv_{df} \exists p(F = [\lambda y\, p])$$

Note that if a property term Π is known to be empty (either as a theorem or by hypothesis), then by (107.1), it follows that $\neg(\Pi = [\lambda y\, p])$, for any p; i.e., it follows that $\neg\exists p(\Pi = [\lambda y\, p])$. So, in that case, we can derive $\neg Propositional(\Pi)$. This shows that no existence clauses are needed in the definiens to ensure that the definiendum is false when the variable F in the definition is instanced by an empty term.

(**276**) **Theorems:** Existence of Propositional Properties. (.1) For any proposition (or state of affairs) p, the propositional property *being an individual such that p* exists.

(.1) $\forall p([\lambda y\, p]\downarrow)$

Next, we leave it to the reader to find two *different* proofs of the following theorem schema, namely, that (.2) *being an individual such that* φ exists:

[184] Thus, in second-order logic with $x = y$ defined as $\forall F(Fx \equiv Fy)$, any time a formula φ (with free variable x) is derivable but φ^y_x is not, then one can establish $x \neq y$.

(.2) $[\lambda\nu\,\varphi]\downarrow$, where φ is any formula in which ν doesn't occur free

Moreover, (.3) if F is *being such that* p, then necessarily, for any x, x exemplifies F iff p is true:

(.3) $F = [\lambda y\, p] \rightarrow \Box\forall x(Fx \equiv p)$

Finally, (.4) if F is propositional, then necessarily F is propositional:

(.4) $\mathit{Propositional}(F) \rightarrow \Box\mathit{Propositional}(F)$

Propositional properties play an extremely important role in some of the applications of object theory in later chapters.

(**277**) **Definition:** Indiscriminate Properties. Let us say that a property F is *indiscriminate* if and only if necessarily, if anything exemplifies F then everything exemplifies F:

$$\mathit{Indiscriminate}(F) \equiv_{df} F\downarrow \,\&\, \Box(\exists xFx \rightarrow \forall xFx)$$

Clearly, then, if F is indiscriminate, then there are no objects x and y such that Fx and $\neg Fy$ (exercise).

Exercise. Explain why $F\downarrow$ is needed as a conjunct in the definiens, i.e., show that without this existence clause, then when Π is an empty property term, one can derive the simpler definiens $\Box(\exists x\Pi x \rightarrow \forall x\Pi x)$, and hence $\mathit{Indiscriminate}(F)$, using axiom (39.5.a) and standard quantifier and modal reasoning.

(**278**) **Theorem:** Propositional Properties are Indiscriminate. It follows from our two previous definitions that propositional properties are indiscriminate:

$$\mathit{Propositional}(F) \rightarrow \mathit{Indiscriminate}(F)$$

(**279**) **Theorem:** Other Facts About Indiscriminate Properties. Some of the following facts will prove useful in later chapters: (.1) necessary properties are indiscriminate; (.2) impossible properties are indiscriminate; (.3) $E!$, $\overline{E!}$, $O!$, and $A!$ are not indiscriminate; and (.4) $E!$, $\overline{E!}$, $O!$, and $A!$ are not propositional properties.

(.1) $\mathit{Necessary}(F) \rightarrow \mathit{Indiscriminate}(F)$

(.2) $\mathit{Impossible}(F) \rightarrow \mathit{Indiscriminate}(F)$

(.3) (.a) $\neg\mathit{Indiscriminate}(E!)$

(.b) $\neg\mathit{Indiscriminate}(\overline{E!})$

(.c) $\neg\mathit{Indiscriminate}(O!)$

(.d) $\neg\mathit{Indiscriminate}(A!)$

(.4) (.a) $\neg Propositional(E!)$

(.b) $\neg Propositional(\overline{E}!)$

(.c) $\neg Propositional(O!)$

(.d) $\neg Propositional(A!)$

(**280**) **Theorems:** Propositional Properties, Necessity, and Possibility. The following claims about propositional properties can be established: (.1) if F might be a propositional property, then it is one; (.2) if F isn't a propositional property, then necessarily it isn't; (.3) if F is a propositional property, then necessarily it is; and (.4) if F might not be a propositional property, then it isn't one:

(.1) $\Diamond\exists p(F = [\lambda y\, p]) \rightarrow \exists p(F = [\lambda y\, p])$

(.2) $\forall p(F \neq [\lambda y\, p]) \rightarrow \Box\forall p(F \neq [\lambda y\, p])$

(.3) $\exists p(F = [\lambda y\, p]) \rightarrow \Box\exists p(F = [\lambda y\, p])$

(.4) $\Diamond\forall p(F \neq [\lambda y\, p]) \rightarrow \forall p(F \neq [\lambda y\, p])$

(**281**) **Theorems:** Propositional Properties and Encoding. It is provable that: (.1) if it is possible that every property that x encodes is propositional, then in fact every property x encodes is propositional, and (.2) if every property that x encodes is propositional, then necessarily every property x encodes is propositional:

(.1) $\Diamond\forall F(xF \rightarrow \exists p(F = [\lambda y\, p])) \rightarrow \forall F(xF \rightarrow \exists p(F = [\lambda y\, p]))$

(.2) $\forall F(xF \rightarrow \exists p(F = [\lambda y\, p])) \rightarrow \Box\forall F(xF \rightarrow \exists p(F = [\lambda y\, p]))$

Intuitively, this is a consequence of the rigidity of encoding; the condition of encoding only propositional properties is rigid (modally collapsed) and doesn't admit of modal distinctions.

9.13 Explanatory Remarks on Definitions

The discussion in the following Remarks help to justify the theory of definitions-by-= developed earlier in this chapter. The first Remark, (282) is designed to help one see why the classical theory of definitions designed for classical logic needs to be refined for the modal contexts in the present system. Remark (282) doesn't require any familiarity with our previous discussion of the inferential role of definitions. The second Remark (283) is designed to help one understand the motivations underlying the primitive Rule of Definition by Identity

(73). The final Remark (284) offers examples of definitions for which the proof that the definiens is significant fails to be modally strict. This Remark may be skipped by those who aren't concerned at present with the application of the system.

(282) **Remark:** Transitioning Away From the Classical Theory of Definitions by Identity – The Problem of Modality. Let's consider the ways in which the classical theory of definitions-by-= fails for a system like the present one. The classical theory of such definitions is usually formulated for a non-modal, first-order predicate calculus with identity, but *without* definite descriptions. In such systems, it is sometimes provable that $\exists!x\varphi$ where, for simplicity, x is the only variable that occurs free in φ. In such situations, one may introduce, by definition, a new individual constant, say δ, to designate the object satisfying φ. But since there are no definite descriptions in the language, one can't formulate the definition as $\delta \ =_{df} \ \imath x\varphi$. Instead, when $\vdash \exists!x\varphi$, one must introduce δ with a definition-by-$\equiv$, i.e., by stipulating:

$$\delta = x \ \equiv_{df} \ \varphi$$

(cf. Suppes 1957, 159–60; Gupta 2014, Section 2.4). For the purposes of our discussion, we shall continue to assume that x is the only variable that occurs free in φ.

Now suppose that we allowed such definitions within the present system. Then since the above example implicitly guarantees that the (closures of the) equivalence $\delta = x \equiv \varphi$ are axioms, it correctly follows that $\delta\downarrow$.[185] Moreover, the definition (also correctly) allows one to eliminate δ from any formula in which it subsequently occurs.[186]

[185] The reasoning is straightforward. The definition licenses the axiom $\forall x(\delta = x \equiv \varphi)$, since this is a universal closure of the equivalence $\delta = x \equiv \varphi$. But, by hypothesis, $\exists!x\varphi$ is a theorem. Then suppose b is such an object, so that we know both φ^b_x and $\forall y(\varphi^y_x \rightarrow y = b)$, by the definition of the uniqueness quantifier (127). Instantiating the universal claim to b, it follows that $\delta = b \equiv \varphi^b_x$. Since we know φ^b_x, it follows that $\delta = b$. So by (107.1), it follows that $\delta\downarrow$.

[186] To see this, we show that ψ^δ_x, i.e., any formula ψ in which δ has been substituted for all the free occurrences of x, is equivalent to $\exists!x\varphi \,\&\, \exists x(\varphi \,\&\, \psi)$. ($\rightarrow$) Assume ψ^δ_x. Since we've just seen that the definition of δ implies $\delta\downarrow$, we can conclude $\delta = \delta \equiv \varphi^\delta_x$ by instantiating δ into $\forall x(\delta = x \equiv \varphi)$ (a universal closure licensed by the definition). And we can independently infer $\delta = \delta$ by instantiating δ into $\forall x(x = x)$, which is obtained by GEN from theorem (117.1). Hence, φ^δ_x. Conjoining what we know, we have $\varphi^\delta_x \,\&\, \psi^\delta_x$. Hence, by $\exists$I, $\exists x(\varphi \,\&\, \psi)$. But, by hypothesis, $\exists!x\varphi$ is a theorem. Hence, $\exists!x\varphi \,\&\, \exists x(\varphi \,\&\, \psi)$.

($\leftarrow$) Assume $\exists!x\varphi \,\&\, \exists x(\varphi \,\&\, \psi)$, to show ψ^δ_x. Let a be a witness to the first conjunct, so that we know:

(ζ) $\varphi^a_x \,\&\, \forall y(\varphi^y_x \rightarrow y = a)$

And let b be a witness to the second conjunct, so that we know:

(ξ) $\varphi^b_x \,\&\, \psi^b_x$

Then by the first conjunct of (ζ) and the definition of δ, it follows that $\delta = a$. It follows from this and the second conjunct of (ζ) that $\forall y(\varphi^y_x \rightarrow y = \delta)$. But this and the first conjunct of (ξ) imply $b = \delta$. So we may substitute δ for b in the second conjunct of (ξ) to conclude ψ^δ_x.

But this method of introducing new individual constants would be disastrous for the present system. To explain this by way of an example, suppose that the following two claims are provable as theorems:

(a) $\exists!x\varphi$

(b) $\Diamond\neg\exists x\varphi$

Since (a) is a theorem, the classical theory would allow one to stipulate:

(c) $\delta = x \ \equiv_{df} \ \varphi$

Definition (c) would, in our system, allow us to take the (closures of) $\delta = x \equiv \varphi$ as axioms. So, for example, we would be able to take the following as axioms:

(d) $\forall x(\delta = x \equiv \varphi)$

(e) $\forall x\Box(\delta = x \equiv \varphi)$

But these would allow us to derive $\Box\exists x\varphi$, which contradicts (b).[187] If well-formed definitions introduce contradictions, then something has gone wrong.

One might suggest here that in order to use (c) as a definition in our system, we have to require more than just $\vdash \exists!x\varphi$. Instead, the suggestion goes, the condition $\vdash \Box\exists!x\varphi$ is required. Such a condition would block the example we've been discussing since (b) couldn't be a theorem if $\vdash \Box\exists!x\varphi$. The latter implies, *a fortiori*, that $\vdash \Box\exists x\varphi$, but if (b) were a theorem, then $\neg\Box\exists x\varphi$ would be as well. Since the example essentially assumes that a contradiction is derivable, it can be discounted.

But this suggestion for preserving the classical theory doesn't work. We can see why, at least intuitively, if we temporarily speak in the familiar idiom of semantically-primitive possible worlds. Suppose it were a theorem that $\Box\exists!x\varphi$. This would be true even if there were just two possible worlds, $\boldsymbol{w}_\alpha$ (the actual world) and $\boldsymbol{w}_1$, and two distinct objects a and b such that a is uniquely φ at $\boldsymbol{w}_\alpha$ and b is uniquely φ at $\boldsymbol{w}_1$. In this modal situation, the equivalence licensed by

[187] To see why, suppose b is some witness to (a), so that we know φ^b_x and $\forall y(\varphi^y_x \rightarrow y = b)$, by the definition of the uniqueness quantifier $\exists!x\varphi$ (127). Then by (d) it follows that:

(f) $\delta = b \equiv \varphi^b_x$

and by (e) it follows that:

(g) $\Box(\delta = b \equiv \varphi^b_x)$

Now by the reasoning in footnote 185, definition (c) implies $\delta\downarrow$. Hence by (125), we know:

(h) $\delta = b \rightarrow \Box\delta = b$

So we can establish $\varphi^b_x \rightarrow \Box\varphi^b_x$ by a hypothetical syllogism chain, as follows: $\varphi^b_x \rightarrow \delta = b$, by (f); $\delta = b \rightarrow \Box\delta = b$, by (h); and $\Box\delta = b \rightarrow \Box\varphi^b_x$, by (g) and the modal theorem $\Box(\psi \equiv \chi) \rightarrow (\Box\psi \equiv \Box\chi)$ (158.6). Having thus established that $\varphi^b_x \rightarrow \Box\varphi^b_x$, then since we know φ^b_x, it follows that $\Box\varphi^b_x$. Hence, by $\exists$I, $\exists x\Box\varphi$. So by the Buridan formula (168.1), $\Box\exists x\varphi$, which contradicts (b).

definition (c), namely $\delta = a \equiv \varphi^a_x$, would fail to be necessary. For since the terms of our language are rigid, the definition would introduce δ as a rigid designator of a, since a is uniquely φ at $\boldsymbol{w}_\alpha$. So $\delta = a$ would be true at $\boldsymbol{w}_1$, since δ rigidly denotes a. But φ^a_x would be false at $\boldsymbol{w}_1$ since, by hypothesis, b is uniquely φ at $\boldsymbol{w}_1$. Hence, the equivalence $\delta = a \equiv \varphi^a_x$ would fail to be true at $\boldsymbol{w}_1$. So the universalized modal equivalence (e), which is licensed by definition (c), can't be true, since it has false instances.

But another suggestion along these lines presents itself, namely, that definition (c) becomes legitimate if we require that $\exists!x\Box\varphi$, instead of $\Box\exists!x\varphi$, be a theorem. Unfortunately, this suggestion fails as well. That's because $\exists!x\Box\varphi$ can be true while $\exists!x\varphi$ is not. Intuitively, from the fact that there is exactly one thing which is φ at every possible world, it doesn't follow that there is exactly one thing which is in fact φ. Suppose there were just two things a and b, and just two worlds $\boldsymbol{w}_\alpha$ and $\boldsymbol{w}_1$, and that a exemplifies P at both $\boldsymbol{w}_\alpha$ and $\boldsymbol{w}_1$, and that b exemplifies P only at $\boldsymbol{w}_\alpha$. Then, in that modal situation, at $\boldsymbol{w}_\alpha$, there is exactly one object (namely a) that exemplifies P at every world, i.e., $\exists!x\Box Px$. But it is not the case, at $\boldsymbol{w}_\alpha$, that there is exactly one thing that exemplifies P, since both a and b exemplify P there. Thus, given only that $\vdash \exists!x\Box\varphi$, we can't *define* δ by saying $\delta = x \equiv_{df} \varphi$, since $\vdash \exists!x\Box\varphi$ doesn't guarantee $\vdash \exists!x\varphi$.[188]

By now, it may be apparent that if one wants to stipulate $\delta = x \equiv_{df} \varphi$, the conditions $\vdash \exists!x\Box\varphi$ and $\vdash \exists!x\varphi$ are both required. But we shall not adapt the classical theory of definitions by introducing new individual constants in this way. Instead, we shall typically introduce a new individual constant when we know that $\vdash \imath x\varphi\downarrow$ and x is the sole variable that occurs free in φ. For then, the definition-by-identity:

(c′) $\delta =_{df} \imath x\varphi$

introduces the rigidly-designating constant δ by way of a significant, rigidly-designating description. This blocks the example that was problematic for the classical theory because (c′) doesn't license the equivalence $\delta = x \equiv \varphi$ as axiomatic. Instead, as we'll see in the next Remark, (c′) will imply the identity $\delta = \imath x\varphi$ when it is known, by proof or by hypothesis, that $\imath x\varphi\downarrow$. And we'll see that one need not require that there be a proof of $\imath x\varphi\downarrow$ to introduce (c′) – the inferential role of definitions by identity will imply that $\neg\delta\downarrow$ if $\neg\imath x\varphi\downarrow$. So we shall defer further discussion of the inferential role of (c′) until Remark (283). Moreover, we postpone, until Remark (284), the discussion of what (c′) implies when the proof that $\imath x\varphi\downarrow$ is a $\star$-theorem.

[188] However, $\vdash \exists!x\Box\varphi$ is sufficient to block the case that led to the original contradiction described above, for one can't simultaneously assert $\exists!x\Box\varphi$ and $\Diamond\neg\exists x\varphi$. A contradiction would ensue without the mediation of any definitions. Moreover, it also blocks the case where, at each world, a *different* witness uniquely satisfies φ, i.e., blocks the case where $\Box\exists!x\varphi$ is true but not $\exists!x\Box\varphi$.

Two final observations are in order. The first is that we should follow the same procedures when introducing a new *relation* constant. We typically introduce such definitions only when the definiens is provably significant. Items (22.1), (22.2), and (230), which contain the definitions of $O!$, $A!$, and $=_E$, respectively, are examples:

$$O! =_{df} [\lambda x \, \Diamond E!x] \tag{22.1}$$

$$A! =_{df} [\lambda x \, \neg\Diamond E!x] \tag{22.2}$$

$$=_E \; =_{df} [\lambda xy \, O!x \,\&\, O!y \,\&\, x = y] \tag{230}$$

These definitions introduce new unary and binary relation *constants* that are provably significant, respectively, by theorems (115.1), (115.2), and the discussion in (229) and (230). In the next Remark we'll see exactly how the above definitions introduce the corresponding identities into our system.

The second and final observation is that the foregoing remarks about the definition of new individual and relation constants have to be generalized to definitions-by-= in which there are free variables in the definiens and definiendum. This brings with it a new set of interesting issues, which are the subject of Remark (283).

(**283**) **Remark:** Definitions by Identity (With Free Variables) and Empty Terms. A definition-by-= takes the following form, in which the distinct variables $\alpha_1, \ldots, \alpha_n$ ($n \geq 0$) occur free:

$$\tau(\alpha_1, \ldots, \alpha_n) =_{df} \sigma(\alpha_1, \ldots, \alpha_n)$$

In any given definition-by-=, the object-language variables that would instance the α_i function as metavariables, by Convention (17.2.a).

The classical understanding of definitions-by-= is that they implicitly assert identity axioms. So where $\tau_1, \ldots, \tau_n$ are any terms substitutable for $\alpha_1, \ldots, \alpha_n$, respectively, and where $\tau(\tau_1, \ldots, \tau_n)$ and $\sigma(\tau_1, \ldots, \tau_n)$ abbreviate, respectively, $\tau^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$ and $\sigma^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$, then on the classical understanding, the above definition would introduce the closures of the following as an axiom schema:

$$\tau(\alpha_1, \ldots, \alpha_n) = \sigma(\alpha_1, \ldots, \alpha_n)$$

But such an understanding won't work in the present system, given the problem of empty terms.

To understand the simplest form of the problem, consider a 'classical' theory such as the theory of real numbers, for illustrative purposes only.[189] In

[189] We discussed this example briefly in footnote 137. In the following discussion, we formulate this definition using a definite description as the definiens, so that we can use a definition-by-= to introduce this kind of division; cf. Suppes 1957, §8.6, which introduces division using a definition-by-≡.

that theory, mathematicians typically use a *conditional* to define division, i.e., to introduce $x/y = z$, where x, y, z range over real numbers. They want to *ignore* terms like $3/0$, $3/(\pi - \pi)$, $3/(3/0)$, etc., and so stipulate that:

> *If* it is a theorem that $y \neq 0$, then the following definition holds,
> $x/y \ =_{df} \imath z(x = y \cdot z)$

But if this definition is supposed to conservatively extend real number theory with new expressions and axioms, this is somewhat awkward. For now, one can't specify the expanded language of real number theory (i.e., the one represented by the definition) so as to only include terms of the form x/y when $y \neq 0$; for any general specification of the language would have to allow terms such as $3/0$, $3/(\pi - \pi)$, $3/(3/0)$, etc., to be well-formed. You can't conditionally specify the language to include those terms only if there is a proof that their denominator is not identical to 0, since you can't specify the proof system until you specify the language. One could perhaps, as a way out, specify a sequence of language and proof system pairs, so that at each pair in the sequence, the proof system of that pair is used to specify the language of the next pair.

But this hardly seems like a good solution to the problem of defining division. What is needed is a way to state the definition in a completely general way but so that (a) it doesn't yield identities in which one of the terms is a division by 0, and (b) it gives one the ability to prove that such facts as $\neg(3/0)\downarrow$, $\neg(3/(\pi - \pi))\downarrow$, $\neg(3/(3/0))\downarrow$, etc. We'll see below how the Rule of Definition by Identity solves this problem. This solution starts by admitting that real number theory is most naturally expressed in a (free) logic that allows for complex terms that fail to have a denotation, such as $3/0$, $3/(\pi - \pi)$, $3/(3/0)$, etc.

So, let's now examine this same problem from the point of view of a theory like the present one, which allows for complex terms that may fail to denote. Consider the following definition, which has a single free variable in the definiens and definiendum and which we introduce for illustrative purposes only (though we've discussed this example in (27) in connection with a different issue):

(ϑ) $\iota_y =_{df} \imath x(x = y)$

This defines ι_y ('the y') as the individual identical to y. We've seen that our system asserts, as a theorem, that the definiens is significant (177.1), and this holds for any value assigned to y. By Convention (17.2), the free variable y in (ϑ) functions as a metavariable, so that we know that (ϑ) has well-formed instances even when a non-denoting individual term is uniformly substituted for y. In other words, (ϑ) is to be interpreted as the following definition schema:[190]

[190] We are ignoring the fact that, strictly speaking, even the variable x should be interpreted as a metavariable, for we now have established that alphabetic variants are interderivable. So the reasons for treating the bound variable x as a metavariable no longer apply.

(ξ) $\iota_\kappa =_{df} \imath x(x = \kappa)$, provided x doesn't occur free in κ

Thus, ι_κ is well-formed, for any individual term κ. When κ is a description of the form $\imath z\varphi$, then the following is an instance of (ϑ), in its interpretation as (ξ):

(ζ) $\iota_{\imath z\varphi} =_{df} \imath x(x = \imath z\varphi)$, provided x doesn't occur free in $\imath z\varphi$

Now, clearly, we can't suppose that the inferential role of (ϑ) or (ξ) is to introduce the axiom schema:

$$\iota_\kappa = \imath x(x = \kappa), \text{ provided } x \text{ doesn't occur free in } \kappa$$

While this would be true when κ is a significant term, it fails when κ is empty; when κ is the provably empty description $\imath z(Pz \,\&\, \neg Pz)$, then the following is not true:

$$\iota_{\imath z(Pz \,\&\, \neg Pz)} = \imath x(x = \imath z(Pz \,\&\, \neg Pz))$$

As we noted in Remark (27), identity statements can't be true when the terms flanking them are empty; nor do we plan to undertake heroic measures such as letting all empty descriptions denote some arbitrary object, such as the null object (265.1). So, what is the inferential role of definitions (ϑ) and (ξ)?

There is an additional issue to keep in mind before we answer this question. Since it is provable that $\neg\imath z(Pz \,\&\, \neg Pz)\downarrow$ and hence that $\neg\imath x(x = \imath z(Pz \,\&\, \neg Pz))\downarrow$, then it should be provable that $\neg\iota_{\imath z(Pz \,\&\, \neg Pz)}\downarrow$. If the definiens is provably empty then the definiendum should be provably empty. After all, the expression $\iota_{\imath z(Pz \,\&\, \neg Pz)}$ is well-formed even if its definiens doesn't have a denotation. So the inferential role of (ϑ) and (ξ) should give us a means to conclude $\neg\iota_{\imath z(Pz \,\&\, \neg Pz)}\downarrow$ when $\neg\imath x(x = \imath z(Pz \,\&\, \neg Pz))\downarrow$. The former is *well-formed* and we know it to be true, but we need a means of proving it.

These observations explain why the classical theory of definition by identity doesn't work for the present system; a new understanding is required. This new understanding is embodied by the primitive Rule of Definition by Identity (73). To simplify the discussion of this metarule, let us consider definitions-by-= with only one free variable, i.e., definitions of the form $\tau(\alpha) =_{df} \sigma(\alpha)$. Then the Rule of Definition by Identity stipulates that a definition of this form introduces an axiom that is a conjunction of two conditionals, expressible in the formal mode as follows: if $\sigma(\tau_1)$ has a denotation, then the equation $\tau(\tau_1) = \sigma(\tau_1)$ is true, and if $\sigma(\tau_1)$ doesn't have a denotation, then neither does $\tau(\tau_1)$. Formally, this would be captured as:

> **Rule of Definition by Identity** (One-Free Variable)
> Whenever τ_1 is substitutable for α in σ, then a definition-by-= of the form $\tau(\alpha) =_{df} \sigma(\alpha)$ introduces the closures of the following, necessary axiom schema:

$$(\omega)\ \ (\sigma(\tau_1)\!\downarrow \to \tau(\tau_1) = \sigma(\tau_1))\ \&\ (\neg\sigma(\tau_1)\!\downarrow \to \neg\tau(\tau_1)\!\downarrow)$$

Though this rule gives rise to some interesting issues (discussed below), let's first see how it addresses the cases discussed so far.

First, consider the definition of division in real number theory. According to the Rule of Definition by Identity, the definition $x/y =_{df} \imath z(x = y \cdot z)$ would introduce the axiom:

$$(\imath z(x = y \cdot z)\!\downarrow \to x/y = \imath z(x = y \cdot z))\ \&\ (\neg\imath z(x = y \cdot z)\!\downarrow \to \neg(x/y)\!\downarrow)$$

So, for $y = 0$, the axiom would assert:

$$(\imath z(x = 0 \cdot z)\!\downarrow \to x/0 = \imath z(x = 0 \cdot z))\ \&\ (\neg\imath z(x = 0 \cdot z)\!\downarrow \to \neg(x/0)\!\downarrow)$$

Since the antecedent of the second conjunct is a theorem of real number theory, it would follow that $\neg(x/0)\!\downarrow$, and so by GEN, $\forall x\neg((x/0)\!\downarrow)$ would be a theorem. Similarly for $y = (\pi - \pi)$. Moreover, when $y = 3/0$, one can show that $\neg(x/(3/0))\!\downarrow$, for any x, since in this case, the rule asserts that the following is axiomatic:

$$(\imath z(x = (3/0) \cdot z)\!\downarrow \to x/(3/0) = \imath z(x = (3/0) \cdot z))\ \&\ (\neg\imath z(x = (3/0) \cdot z)\!\downarrow \to \neg(x/(3/0))\!\downarrow)$$

Since $\neg(3/0)\!\downarrow$, $(3/0) \cdot z$ is provably empty, for any z.[191] So $x = (3/0) \cdot z$ is always false, for any x, implying thereby that $\neg\imath z(x = (3/0) \cdot z)\!\downarrow$. Thus, the rule yields $\neg(x/(3/0))\!\downarrow$, for any x. So the Rule of Definition by Identity handles the definition of division in real number theory in a general way – it extends the language with new terms of the form κ/κ', for arbitrary individual terms κ and κ'; it asserts that the definition yields identities when the definiens is significant; and it allows us to prove that the definiendum is empty when the definiens is empty.

Second, reconsider (ϑ). Let κ be any individual term substitutable for y in $\imath x(x = y)$. Then, on the proposed inferential role, (ϑ) would introduce the following, necessary axiom schema:

$$(\omega')\ \ (\imath x(x = \kappa)\!\downarrow \to \iota_\kappa = \imath x(x = \kappa))\ \&\ (\neg(\imath x(x = \kappa)\!\downarrow) \to \neg\iota_\kappa\!\downarrow).$$

Let's consider how this applies to specific instances, starting with a case where the first conjunct of (ω') is operative, i.e., a case where we can show that the antecedent of the first conjunct of (ω') holds. Suppose κ is the constant a, so that the relevant instance of definition (ϑ) is:

[191] For reductio, suppose $((3/0) \cdot z))\!\downarrow$. Then by the axioms for multiplication, $(3/0) \cdot z = 3z/0$. But by the previous case we examined in the text, we saw that $\neg(x/0)\!\downarrow$, for any x, and so $\neg(3z/0)\!\downarrow$. Hence, $\neg((3/0) \cdot z = 3z/0)$, by theorem (107.2). Contradiction. Thus, $\neg((3/0) \cdot z)\!\downarrow$. Alternatively, just assume that in the formulation of real number theory under consideration, which allows for empty complex terms, multiplication is axiomatized in such a way that the term $\kappa \cdot \kappa'$ is empty if either κ or κ' is empty.

$$\iota_a =_{df} \imath x(x=a)$$

So, the Rule of Definition by Identity says that the inferential role of this instance is to introduce the following necessary axiom (and hence a claim derivable from any premises):

$$(\omega'') \quad (\imath x(x=a)\downarrow \rightarrow \iota_a = \imath x(x=a)) \,\&\, (\neg(\imath x(x=a)\downarrow) \rightarrow \neg\iota_a\downarrow)$$

Clearly, the antecedent of the first conjunct is derivable: since $\imath x(x=y)\downarrow$ is a theorem (177), it follows that $\forall y(\imath x(x=y)\downarrow)$, and since $a\downarrow$ is axiomatic (39.2), it follows that $\imath x(x=a)\downarrow$. So the first conjunct of (ω'') yields $\iota_a = \imath x(x=a)$ as a theorem. On this understanding, (ϑ) enables an appeal to this identity in any reasoning we might conduct.

Now consider a case where the second conjunct of (ϑ') is the operative one. Suppose κ is the description $\imath z\psi$, where ψ is $Pz \,\&\, \neg Pz$, so that the relevant instance of definition (ϑ) is:

$$\iota_{\imath z\psi} =_{df} \imath x(x=\imath z\psi)$$

Given the inferential role stipulated by the Rule of Definition by Identity, the above introduces the following as a necessary axiom:

$$(\omega''') \quad (\imath x(x=\imath z\psi)\downarrow \rightarrow \iota_{\imath z\psi} = \imath x(x=\imath z\psi)) \,\&\, (\neg(\imath x(x=\imath z\psi)\downarrow) \rightarrow \neg\iota_{\imath z\psi}\downarrow)$$

Now the laws governing descriptions will guarantee that $\neg(\imath z\psi\downarrow)$, and so by reasoning described in footnote 102, we can infer $\neg(\imath x(x=\imath z\psi)\downarrow)$. It follows from the second conjunct of (ω''') that $\neg(\iota_{\imath z\psi}\downarrow)$. So on this understanding of (ϑ), we have a proof of $\neg(\iota_{\imath z\psi}\downarrow)$ from the fact that the definiens of $\iota_{\imath z\psi}$ doesn't have a denotation. We may therefore continue any reasoning we're engaged in secure in the knowledge that no true exemplification, encoding, or identity formula will have $\iota_{\imath z\psi}$ as a primary term.

At this point, we have sufficiently motivated the Rule of Definition by Identity (73) as a primitive metarule. The key to its formulation and generalization concerns the issues that arise for the classical understanding of definitions-by-= with free variables in the definiens and definiendum. In (73), the rule is formulated so that it governs definitions with n free variables in the definiens and definiendum, i.e., so that it governs the introduction of n-ary term-forming operators. Thus, the foregoing remarks indicate the conditions under which we can preserve a classical understanding of the inferential role of term-forming operators. We can rest assured that term-forming operators introduced into our language by definition are logically well-behaved if they are significant when applied to appropriate arguments.

There is, however, one final issue to discuss. By guaranteeing that the definiendum is empty whenever the definiens is empty, the Rule of Definition by Identity does implement the *garbage in*, *garbage out* principle, but only in the first instance. For the rule still allows one to introduce new, *impractical* terms that have denotations even though the arguments to the term-forming operator do not. To illustrate the issue, we'll consider some examples of new terms of the form $\tau(\tau_1)$ that have denotations even though the argument term τ_1 does not. But the reader should note: these examples will not lead us to further refine the Rule of Definition by Identity so as to ensure that $\tau(\tau_1)$ is provably empty if τ_1 is empty. That's because the system already has terms τ that denote even though τ contains empty subterms, and so adding new, defined terms of this sort won't introduce logical problems. Moreover, the refinements needed to forestall new, impractial terms would not only complicate the Rule of Definition by Identity significantly, but would also give rise to a problem of its own that requires even further revisions. So the following discussion is primarily for edification purposes; those interested in the issues surrounding a more sensitive rule might find the following discussion to be of value.

To see how the Rule of Definition by Identity yields *impractical* terms, consider the following definitions of the individual-term-forming operator $\boldsymbol{a}_{(\,)}$ and the relation-term-forming operator $\overline{(\,)}$:

(A) $\boldsymbol{a}_y \ =_{df} \ \imath x(A!x \,\&\, \forall F(xF \equiv Fy))$

(B) $\overline{F} \ =_{df} \ [\lambda y\, \neg Fy]$

Intuitively, the intent of these definitions is to identify new individuals and relations in terms of some given individuals and relations. But, in object theory, we shall be able to prove the following:

(a) $\imath x(A!x \,\&\, \forall F(xF \equiv F\kappa))\downarrow$, for *any* individual term κ in which x and F don't occur free

(b) $[\lambda y\, \neg\Pi y]\downarrow$, for *any* property term Π in which y doesn't occur free

It is worth a brief digression to explain why these hold even when κ and Π are terms that fail to have a denotation, though (b) should already be familiar from the discussion of example (E) in Remark (155).

If κ is a *non-denoting* description in which x and F don't occur free, say $\imath z\psi$, then consider the canonical description $\imath x(A!x \& \forall F(xF \equiv F\imath z\psi))$. This canonical description has a denotation; it denotes the abstract object that encodes no properties, since in such a case no property F is such that $F\imath z\psi$. It is a theorem that there is a unique abstract object that encodes no properties (264). So (a) is true when κ is assigned the non-denoting description $\imath z\psi$. The current Rule of Definition by Identity allows us to derive, from definition (A), that $\boldsymbol{a}_{\imath z\psi} = \imath x(A!x \,\&\, \forall F(xF \equiv F\imath z\psi))$. Thus, the new term $\boldsymbol{a}_{\imath z\psi}$ has a denotation even though its argument $\imath z\psi$ does not.

Similarly, if Π is a *non-denoting* property expression in which y doesn't occur free, say $[\lambda z\, \psi]$, then consider the λ-expression $[\lambda y\, \neg[\lambda z\, \psi]y]$. This expression has a denotation; it denotes a universal property. Intuitively, if $[\lambda z\, \psi]$ denotes nothing, then $[\lambda z\, \psi]y$ is (universally) false and so $\neg[\lambda z\, \psi]y$ is (universally) true. So, in such a case, $[\lambda y\, \neg[\lambda z\, \psi]y]$ denote a property that everything exemplifies. Therefore (b) is true when Π is $[\lambda z\, \psi]$. The current Rule of Definition by Identity implies, given definition (B), that $\overline{[\lambda z\, \psi]} = [\lambda y\, \neg[\lambda z\, \psi]y]$. In this case, $\overline{[\lambda z\, \psi]}$ has a denotation even though $[\lambda z\, \psi]$ does not – recall Remark (198).

So we've now seen:

- examples of empty terms κ and Π that occur as arguments in provably significant terms of the form (a) and (b),
- that, in these cases, the inferential role of (A) and (B) assigned by the Rule of Definition by Identity allows us to derive an identity claim governing $\boldsymbol{a}_\kappa$ and $\overline{\Pi}$, and
- that, in these cases, the definienda thereby provably have a denotation even though they consist of a term-forming operator that is operating on a non-denoting argument.

Since there are no significant logical issues posed by these facts, and the complications needed to forestall them are significant, we may simply ignore them. But if one were to rigorously pursue the *garbage in*, *garbage out* principle, one could take the following steps.

One could refine the Rule of Definition by Identity by having it stipulate: if the arguments to the definiens all have a denotation and the definiens (with the arguments substituted for the free variables) has a denotation, then the equation between the definiendum and definiens holds; otherwise, the definiendum fails to denote. In the case of a definition with one free variable, this could be stated formally as:

Alternative Rule of Definition by Identity (1 free variable)
Whenever τ_1 is substitutable for α in σ, then a definition-by-= of the form $\tau(\alpha) =_{df} \sigma(\alpha)$ introduces the closures of the following, necessary axiom schema:

$$(\omega)\;\; ((\tau_1{\downarrow} \,\&\, \sigma(\tau_1){\downarrow}) \rightarrow (\tau(\tau_1) = \sigma(\tau_1))) \,\&\, ((\neg\tau_1{\downarrow} \vee \neg\sigma(\tau_1){\downarrow}) \rightarrow \neg\tau(\tau_1){\downarrow})$$

Before we explore how this would imply that impractical terms are empty, note that how this rule would be generalized for the definition of term-forming operators that take n variables ($n \geq 0$):

Alternative Rule of Definition by Identity (n free variables)
Whenever $\tau_1,\ldots,\tau_n$ are substitutable, respectively, for $\alpha_1,\ldots,\alpha_n$ in σ, then a definition-by-= of the form $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$ introduces the closures of the following, necessary axiom schema:

$$(\omega)\ \ [(\tau_1\!\downarrow \,\&\, \ldots \,\&\, \tau_n\!\downarrow \,\&\, \sigma(\tau_1,\ldots,\tau_n)\!\downarrow) \to \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n)]\ \& \\ [(\neg\tau_1\!\downarrow \vee \ldots \vee \neg\tau_n\!\downarrow \vee \neg\sigma(\tau_1,\ldots,\tau_n)\!\downarrow) \to \neg\tau(\tau_1,\ldots,\tau_n)\!\downarrow]$$

The axiom (ω) in this rule *stipulates* that the identity $\tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n)$ holds if all of the arguments $\tau_1,\ldots,\tau_n$ and the definiens $\sigma(\tau_1,\ldots,\tau_n)$ are significant *and* that the new term $\tau(\tau_1,\ldots,\tau_n)$ is empty if any of the arguments $\tau_1,\ldots,\tau_n$ or the definiens $\sigma(\tau_1,\ldots,\tau_n)$ are empty.

Note that the Alternative Rule of Definition by Identity still correctly allows one to derive the following identities from (A) and (B), respectively:

(i) $\boldsymbol{a}_y = \imath x(A!x \,\&\, \forall F(xF \equiv Fy))$

(ii) $\overline{F} = [\lambda y\, \neg Fy]$

In each case, the first conjunct of the Alternative Rule is operative, since not only is it axiomatic that $y\!\downarrow$ and $F\!\downarrow$, but the definiens in each of (A) and (B) is provably significant.

But, more significantly, definitions (A) and (B), under the Alternative Rule, would guarantee, respectively, that:

- $\neg\boldsymbol{a}_{\imath z\psi}\!\downarrow$ is a theorem whenever $\neg\imath z\psi\!\downarrow$ is a theorem
- $\neg\overline{[\lambda z\, \psi]}\!\downarrow$ is a theorem whenever $\neg[\lambda z\, \psi]\!\downarrow$ is a theorem

These results follow from the second conjunct of the axiom schema implicitly introduced by the definitions. So, given this alternative inferential role, to show that the definiendum fails to denote, it suffices to show that the argument to the term-forming operator fails to denote.

But even this Alternative Rule of Definition by Identity has to be further refined. A problem arises in connection with the definition of new 0-ary relation terms with free variables. The problem is that the Alternative Rule is inconsistent with the theorem that every 0-ary relation term Π^0 is significant. Both $\Pi^0\!\downarrow$ (104.1) and $\varphi\!\downarrow$ (104.2) are theorems, for any 0-ary relation term Π^0 or formula φ. But suppose we're introducing, by definition, a new 0-ary term-forming operator $\Pi(\ )$ and suppose further that it takes a free individual term as an argument, so that the definition takes the form $\Pi(x) =_{df} \sigma(x)$, where $\sigma(x)$ is a 0-ary relation term. Then, when the argument is a description, say $\imath z\psi$, the Alternative Rule of Definition by Identity stipulates that the following is an axiom:

$$((\imath z\psi\!\downarrow \,\&\, \sigma(\imath z\psi)\!\downarrow) \to (\Pi(\imath z\psi) = \sigma(\imath z\psi)))\ \&\ ((\neg\imath z\psi\!\downarrow \vee \neg\sigma(\imath z\psi)\!\downarrow) \to \neg(\Pi(\imath z\psi)\!\downarrow))$$

So if $\imath z\psi$ is $\imath z(Pz \,\&\, \neg Pz)$ and $\neg\imath z\psi\downarrow$ is thereby derivable, the second conjunct is operative, thereby implying $\neg(\Pi(\imath z\psi)\downarrow)$. But this contradicts $\Pi(\imath z\psi)\downarrow$, which is an instance of theorem (104.2).

To address this new problem, one might propose:

Second Alternative Rule of Definition by Identity
If $\tau_1,\ldots,\tau_n$ are substitutable for $\alpha_1,\ldots,\alpha_n$, respectively, in σ, then a definition-by-= of the form $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$ introduces the closures of the following, necessary axiom schema:

(ω) $[(\tau_1\downarrow \,\&\, \ldots \,\&\, \tau_n\downarrow \,\&\, \sigma(\tau_1,\ldots,\tau_n)\downarrow) \rightarrow \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n)] \,\&$
$[(\neg\tau_1\downarrow \vee \ldots \vee \neg\tau_n\downarrow \vee \neg\sigma(\tau_1,\ldots,\tau_n)\downarrow) \rightarrow \neg\tau(\tau_1,\ldots,\tau_n)\downarrow]$,
whenever $\sigma(\alpha_1,\ldots,\sigma_n)$ is an individual term, an m-ary relation term ($m \geq 1$), or a 0-ary relation term in which none of $\alpha_1,\ldots,\alpha_n$ are individual variables or k-ary relation variable for $k \geq 1$,

and:

(ω') $[(\tau_1\downarrow \,\&\, \ldots \,\&\, \tau_n\downarrow \,\&\, \sigma(\tau_1,\ldots,\tau_n)\downarrow) \rightarrow \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n)] \,\&$
$[(\neg\tau_1\downarrow \vee \ldots \vee \neg\tau_n\downarrow \vee \neg\sigma(\tau_1,\ldots,\tau_n)\downarrow) \rightarrow \Box\neg\tau(\tau_1,\ldots,\tau_n)]$,
whenever $n \geq 1$ and $\sigma(\alpha_1,\ldots,\alpha_n)$ is a 0-ary relation term in which one or more of $\alpha_1,\ldots,\alpha_n$ is an individual variable or k-ary relation variable for $k \geq 1$.

Though this lays the problem of defining new 0-ary term-forming operators to rest, the reader might, at this point, be convinced that the official version of the rule suffices and stands in need of no further refinement. The cost of ruling out impractical terms is significant and since there is no real harm in living with impractical terms that arise by definition, we have chosen not to pursue the above alternatives.

(**284**) **Remark:** When the Significance of the Definiens Rests on a $\star$-Theorem. Consider, again, the general case of any individual constant δ introduced by the following definition, in which $\imath x\varphi$ is a closed term:

$$\delta =_{df} \imath x\varphi$$

The inferential role of this definition, as we just saw in Remark (283) is to introduce the following as a necessary axiom:

(ω) $(\imath x\varphi\downarrow \rightarrow (\delta = \imath x\varphi)) \,\&\, (\neg\imath x\varphi\downarrow \rightarrow \neg\delta\downarrow)$

Now suppose that we can prove $\imath x\varphi\downarrow$, but that the proof rests on a $\star$-theorem, so that $\imath x\varphi\downarrow$ is a $\star$-theorem. Then when we appeal to $\imath x\varphi\downarrow$ to detach the consequent from the antecedent of the first conjunct of (ω), the resulting identity $\delta = \imath x\varphi$ becomes a $\star$-theorem. And since identities are necessary, $\Box(\delta = \imath x\varphi)$

becomes a $\star$-theorem. This is as it should be: a necessary identity has been proved as a theorem on the basis of a contingency and should therefore be flagged as such.

To see how this works in practice, suppose one were to extend our theory by adding, as an axiom, the contingent claim that there exists a unique moon of Earth. If we use 'e' as the name of Earth and represent this axiom as $\exists!xMxe$, then we would annotate the axiom as *modally fragile* and decorate its item number with a $\star$.[192] Since $\exists!xMxe$ is asserted as a modally fragile axiom, all of its $\Box$-free closures are also taken to be modally fragile axioms and so become $\star$-theorems. So by theorem (144.1)$\star$, or by RA and theorem (176.2), it would follow that $\imath xMxe\downarrow$ is a $\star$-theorem. Since $\imath xMxe$ is provably significant, we might introduce a name, say m, to designate this unique object, via the following definition:

$$m =_{df} \imath xMxe$$

Given the inferential role of definitions-by-= described in (283), this introduces the necessary axiom:

$$(\omega_1)\quad (\imath xMxe\downarrow \to (m = \imath xMxe))\ \&\ (\neg\imath xMxe\downarrow \to \neg(m\downarrow))$$

In this scenario, note that all of the following become provable as $\star$-theorems:

- $\star$ $m = \imath xMxe$, by (ω_1) and the $\star$-theorem $\imath xMxe\downarrow$
- $\star$ $m\downarrow$, by the $\star$-theorem $m = \imath xMxe$ and the $\Box$-theorem (107.1)
- $\star$ $\Box\imath xMxe\downarrow$, by the $\star$-theorem $\imath xMxe\downarrow$ and the $\Box$-theorem (106)
- $\star$ $\Box m\downarrow$, by the $\star$-theorem $m\downarrow$ and the $\Box$-theorem (106)
- $\star$ $\Box m = \imath xMxe$, by the $\star$-theorem $m = \imath xMxe$ and the $\Box$-theorem (125.2)

And so on. These consequences, again, are as they should be: all are claims *derived from an axiom whose necessitation was not asserted* and so flagged as such.[193]

We conclude these observations by generalizing the discussion to relation terms. Suppose $[\lambda x_1 \ldots x_n\, \varphi]$ is a closed term and that the proof of $[\lambda x_1 \ldots x_n\, \varphi]\downarrow$ is a $\star$-theorem. If we introduce a new n-ary relation constant, say P_{10} by the definition:

[192] Since the claim $\exists!xMxe$ is, intuitively, contingent, we could also assert $\Diamond\neg\exists!xMxe$ as a (necessary) axiom. But the possibility claim is not needed for the present discussion; it suffices that the axiom is marked as modally fragile, for that just means, by definition (42), that we're not asserting its modal closures as axioms.

[193] See Remark (695) in Chapter 13, Section 13.3, where Leibniz's notion of a *hypothetical necessity*, which he deploys in defense of his containment theory of truth against an objection by Arnauld, is analyzed as a necessary truth derived on the basis of a contingency.

$$P_{10} =_{df} [\lambda x_1 \ldots x_n\, \varphi]$$

then the identity claim $P_{10} = [\lambda x_1 \ldots x_n\, \varphi]$ becomes a $\star$-theorem, given the inferential role of a definition-by-= described in (283), for the reasons just discussed in connection with new individual constants. The reader might try to construct an example of a λ-expression whose significance is provable only by appeal to a contingency.[194]

[194]The λ-expression $[\lambda y\, y =_E \imath xMxe]$, when $\exists!xMxe$ has been asserted as a modally fragile axiom, is *not* such an example. The resulting $\star$-theorem, $\imath xMxe\downarrow$, isn't required to establish the modally strict claim that $[\lambda y\, y =_E \imath xMxe]\downarrow$, since (a) the λ-expression is a core λ-expression and so this claim axiomatic by (39.2) or (b) the claim is provable in the manner discussed at the end of Remark (231). So it takes some work to develop an example of a λ-expression $[\lambda x\, \varphi]$ such that $[\lambda x\, \varphi]\downarrow$ is a $\star$-theorem.

Here is one suggestion: let $\imath xMxe\downarrow$ be the $\star$-theorem from the above example and consider the λ-expression $[\lambda x\, \imath xMxe\downarrow]$. It can be shown that $[\lambda x\, \imath xMxe\downarrow]\downarrow$ is a $\star$-theorem. In the discussion above we established that $\Box \imath xMxe\downarrow$ is a $\star$-theorem. From this, and the fact that $\Box(E!x \to E!x)$, it follows by (158.7) that $\Box((E!x \to E!x) \equiv \imath xMxe\downarrow)$ is a $\star$-theorem. So by GEN and then an application of BF (167.1), it follows that $\Box\forall x((E!x \to E!x) \equiv \imath xMxe\downarrow)$ is a $\star$-theorem. So by axiom (49) and the fact that $[\lambda x\, E!x \to E!x]\downarrow$, it follows that $[\lambda x\, \imath xMxe\downarrow]\downarrow$ is a $\star$-theorem.

Chapter 10

Basic Logical Objects

In this chapter, we prove the existence of some basic logical objects and some fundamental theorems about them. Such objects include the truth-value of proposition p ($\circ p$), the extension of property F (ϵF), the class of Fs ($\{y|Fy\}$), the direction of line a ($\vec{a}$), the shape of figure c ($\tilde{c}$), etc. We also generalize these applications to develop theorems governing any logical object abstracted from an equivalence condition or an equivalence relation.[195]

10.1 Truth-Values

(**285**) **Remark:** On Truth-Values. Frege postulated truth-values in his lecture of 1891 (13), and they are the very first logical objects that he officially introduces in his *Grundgesetze der Arithmetik*; they appear in Volume I, §2, just after the section on functions (see Frege 1893, 7). In what follows, we identify truth-values as abstract, logical objects, prove they exist, and further prove, among other things: (a) that necessarily there are exactly two truth-values, and (b) that the truth-value of p is identical to the truth-value of q if and only if p is equivalent to q. The theorems about truth-values proved below are often principles that Frege implicitly assumed. The key idea underlying truth-values of

[195]Some of the definitions and theorems below are revised and enhanced versions of those developed in Anderson & Zalta 2004. The main difference is that, in the present work, we develop definitions of such notions as *truth-value*, *class*, etc., so as to allow for modally strict theorems about these notions. By contrast, in Anderson & Zalta 2004, the definitions often used rigid definite descriptions, thereby making it difficult to establish modally strict theorems. Compare, for example, the definition of x is a *truth-value* in Anderson & Zalta (2004, 14) with the one developed below in (290). There are modally strict facts about truth-values proved below for which the version in Anderson & Zalta 2004 not modally strict. For example, theorem (292), that there are exactly two truth-values, is modally strict in the present work, but the corresponding version in Anderson & Zalta (2004, 14) is not.

propositions is that, given a proposition p, the condition, *q is materially equivalent to p*, intuitively defines a logical pattern relative to p. This pattern can be objectified into an abstract, logical object that encodes just those properties F having the form $[\lambda y\, q]$ for propositions q materially equivalent to p. Such a logical object is identified as the truth-value of p.

(**286**) **Definitions:** Truth-Value of a Proposition. Recall that in (275) we defined a propositional property to be any property F such that $\exists q(F = [\lambda y\, q])$. Employing our conventions for definitions-by-$\equiv$ in (17), we now say that *x is a truth-value of p* just in case x is an abstract object that encodes all and only those properties F such that for some proposition q materially equivalent to p, F is the propositional property $[\lambda y\, q]$:

$$TruthValueOf(x,p) \equiv_{df} A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q]))$$

If we informally say *F is constructed from q* whenever $F = [\lambda y\, q]$, then we may read the above definition as follows: x is a truth-value of p whenever x is an abstract object that encodes all and only properties constructed from propositions materially equivalent to p.[196]

(**287**) **Remark:** Reminder About Definitions. Note the following facts about the foregoing definition:

- We no longer need to interpret the bound variables under Convention (17.2.b); that is, we need not suppose that the bound variables F, q, and y function as metavariables. By Rule $\equiv$Df, the definition introduces a biconditional axiom. Since axioms are theorems, we can derive, as a theorem, any alphabetic variant of the axiom, by the Rule of Alphabetic Variants (114). Indeed, given Rules GEN, RN, and RA, any closure of the biconditional axiom is derivable as a theorem and, therefore, so are their alphabetic variants. So the discussion in Remarks (32) and (28) no longer applies.

- We no longer need to interpret the free variable p under Convention (17.2.a); that is, we need not suppose that p functions as a metavariable. Since every 0-ary relation term provably has a denotation (104.1), every 0-ary relation term in which F doesn't occur free can be substituted for $\forall p$ in the universally quantified biconditional that can be derived from the definition. So the discussion in Remark (27) no longer applies to the variable p.

[196] We'll discover in Chapter 12, in Section 12.1 on situations, that the above definition can be 'simplified'. But this simplification requires a lot of machinery that isn't strictly necessary for defining, and proving facts about, the notion of a truth-value. See the Exercise at the end of (486).

- The variable x *does* still function as a metavariable, under Convention (17.2.a); every individual term κ in which F doesn't occur free, whether significant or not, can be used to form an instance of the definition in which κ is uniformly substituted for x. So the discussion in Remark (27) still applies to the variable x.

- For instances of the definition involving individual terms κ such that $\vdash \neg\kappa\downarrow$, we don't need to add an existence clause to the definiens to ensure that $\neg \mathit{TruthValueOf}(x,\kappa)$ is provable. If $\vdash \neg\kappa\downarrow$, then $A!\kappa$ is provably false and the definiens will be provably false, thereby yielding a derivation of $\neg \mathit{TruthValueOf}(x,\kappa)$. See the discussion in Remark (36).

See the Remarks indicated if these facts aren't clear.

(**288**) **Theorem:** There Exists a (Unique) Truth-Value of p.

(.1) $\exists x \mathit{TruthValueOf}(x,p)$

(.2) $\exists! x \mathit{TruthValueOf}(x,p)$

So, by applying GEN to (.2), we have established the principle have every proposition has a unique truth-value. Thus $\exists! x \mathit{TruthValueOf}(x,\varphi)$ is a theorem schema provable for any formula φ. Finally, we derive, by modally strict means, an important, unheralded principle, namely (.3) if x is a truth-value of p and y is a truth-value of q, then x is identical to y if and only if p is materially equivalent to q:

(.3) $(\mathit{TruthValueOf}(x,p) \,\&\, \mathit{TruthValueOf}(y,q)) \rightarrow (x = y \;\equiv\; (p \equiv q))$

(**289**) **Lemmas:** Facts About Propositional Properties and Truths (or Falsehoods). Some basic facts about propositional properties are: (.1) p is true if and only if the propositional properties constructed from true propositions are precisely the propositional properties constructed from propositions materially equivalent to p; and (.2) it is not the case that p if and only if the propositional properties constructed from false propositions are precisely the propositional properties constructed from propositions materially equivalent to p):

(.1) $p \equiv \forall F[\exists q(q \,\&\, F = [\lambda y\, q]) \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])]$

(.2) $\neg p \equiv \forall F[\exists q(\neg q \,\&\, F = [\lambda y\, q]) \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])]$

Note that these theorems are modally-strict.

(**290**) **Definition:** Truth-Values. When Frege introduced truth-values in 1891, he not only stipulated that they exist (without defining them) but also asserted that there are exactly two of them, naming them The True and The False. By contrast, our procedure has been to first define what it is for an object to be a *truth-value-of* a proposition and show that for any proposition p, there is a unique such entity. We now define what it is for an object to be a *truth-value* in terms of the notion *truth-value-of*. We say that x is a *truth-value* iff x is a truth-value of some proposition:

$$\mathit{TruthValue}(x) \equiv_{df} \exists p(\mathit{TruthValueOf}(x,p))$$

We now prove some general facts about truth-values and then prove that there are *exactly* two of them. Then we'll define The True and The False as abstract objects and prove that The True and The False are truth-values.

(**291**) **Lemmas:** Abstract Objects That Encode Just The Truths (or Just The Falsehoods) Are Truth-Values. We now prove (.1) if x is an abstract object that encodes all and only properties constructed from true propositions, then it is a truth-value, and (.2) if x is an abstract object that encodes all and only properties constructed from false propositions, then it is a truth-value:

(.1) $(A!x \,\&\, \forall F(xF \equiv \exists q(q \,\&\, F = [\lambda y\, q]))) \rightarrow \mathit{TruthValue}(x)$

(.2) $(A!x \,\&\, \forall F(xF \equiv \exists q(\neg q \,\&\, F = [\lambda y\, q]))) \rightarrow \mathit{TruthValue}(x)$

These facts are modally strict.

(**292**) **Theorem:** There are Exactly Two Truth-Values. We prove the claim that there are exactly two truth-values in the following form:

$$\exists x \exists y[\mathit{TruthValue}(x) \,\&\, \mathit{TruthValue}(y) \,\&\, x \neq y \,\&\, \forall z(\mathit{TruthValue}(z) \rightarrow z = x \vee z = y)]$$

By RN, this is a necessary truth.

(**293**) **Theorem:** The Truth-Value of p Exists. It now follows that *the truth-value of* p exists:

$$\imath x \mathit{TruthValueOf}(x,p)\downarrow$$

So, by GEN, for every proposition p, the truth-value of p exists.

Note that the above theorem can be easily proved in one of two ways: (a) from the previous theorem (288.2), the Rule of Actualization (RA), and a lemma for actuality and existence (176.2), or (b) from an instance of theorem (252), definition (286), and theorem (149.3), by the substitution of identicals.

(**294**) **Definition:** Notation for the Truth-Value of p. The previous theorem guarantees that the description $\imath x \mathit{TruthValueOf}(x,p)$ is significant and so holds for every proposition p. Let us therefore introduce the notation $\circ p$ to refer to p's truth-value:

$$\circ p =_{df} \imath x \mathit{TruthValueOf}(x,p)$$

This definition is interesting in the following respect. We know by (104.1) that $\Pi\downarrow$, for every 0-ary relation term Π. We can therefore prove from (293) that $\imath x TruthValueOf(x,\Pi)\downarrow$, for any 0-ary relation term Π in which x doesn't occur free. And if x does occur free in Π, we can always pick some individual variable that doesn't occur free in Π, say y, to prove $\imath y TruthValueOf(y,\Pi)\downarrow$. So given the inferential role of definitions-by-=, as described in (73), we can establish an identity of the form $\circ\Pi = \imath \nu TruthValueOf(\nu,\Pi)$, for any 0-ary relation term Π. It is interesting when a term-forming operator yields an identity relative to *every* term that can serve as its argument.

(**295**) **Definition:** When Objects Encode Propositions. We now extend the notion of encoding. We say that object x *encodes* proposition p (written $x\Sigma p$) just in case x encodes *being-a-y-such-that-p*, i.e., encodes $[\lambda y\, p]$:

$$x\Sigma p \equiv_{df} x[\lambda y\, p]$$

It is important to remember that although every formula φ signifies a proposition (104.1), one may *not* substitute a formula φ for p in this definition if y occurs free in φ. That would not result in a valid instance of the definition, for the free variable in φ would get captured by λy in the term $[\lambda y\, \varphi]$! As we noted in the Convention for Variables in Definitions (17.2.b), the y in the above definition is functioning as a metavariable. If we state the definition using a metavariable, say ν, instead of y, then we must add the proviso that ν is not free in φ. Clearly, the operator λy in the term $[\lambda y\, p]$ doesn't capture any variable in the matrix p, and so any well-formed instance of the definition has to conform to this standard – the variable bound by the λ may not capture any variable free in the substitution instance for p. So if y does occur free in φ, then we find a variable not free in φ, say z, and regard the definition as asserting $x\Sigma\varphi \equiv_{df} x[\lambda z\, \varphi]$. Thus, the λ-expression in the definiens of any substitution instance of the definition of $x\Sigma p$ will always be a core λ-expression (9.2) and hence significant (39.2).

We henceforth adopt the convention that '$x\Sigma\ldots$' is to be interpreted with the smallest scope possible. For example, $x\Sigma p \to p$ is to be parsed as $(x\Sigma p) \to p$ rather than as $x\Sigma(p \to p)$.

(**296**) **Theorem:** The Truth-Value of p is Canonical. It is an easy consequence of our definitions that (.1) the truth-value of p is identical to the abstract object that encodes exactly the properties F constructed out of propositions materially equivalent to p:

(.1) $\circ p = \imath x(A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])))$

Thus, by (253), $\circ p$ is (identical to) a canonical individual. Moreover, it follows by modally strict means that (.2) the truth-value of p encodes p:

(.2) $\circ p \Sigma p$

(**297**) **Theorem:** When the Submatrix is Contingent. To help set up the observation that $\circ p$ is not strictly canonical, it will prove useful to show that: for some proposition p and some property F, it is possible that both (a) for some proposition q materially equivalent to p, F is the property *being such that* q and (b) it is possible that there is no proposition q materially equivalent to p such that F is *being such that* q:

$$\exists p \exists F \Diamond(\exists q((q \equiv p) \,\&\, F = [\lambda y\, q]) \,\&\, \Diamond\neg\exists q((q \equiv p) \,\&\, F = [\lambda y\, q]))$$

Note that if we let φ be the formula $\exists q((q \equiv p) \& F = [\lambda y\, q])$, then this theorem establishes, for some proposition p and property F, that $\Diamond(\varphi \,\&\, \Diamond\neg\varphi)$. By (165.11), we have established, for some p and F, that $\Diamond\varphi \,\&\, \Diamond\neg\varphi$. Thus, for some values of the free variables, the submatrix of the canonical description of $\circ p$ (296) is a contingent condition. This is a key to the observation that $\circ p$ is not strictly canonical, to which we now turn.

(**298**) **Remark:** $\circ p$ is Not Strictly Canonical. Though we know, by (296) that $\circ p$ is (identical to) a canonical object, we are now in a position to show that it is *not* (identical to) a strictly canonical object. In (260.2), we stipulated that a canonical description $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$ is *strictly* canonical just in case φ is a rigid condition on properties, i.e., by (260.1), just in case $\vdash_\Box \forall F(\varphi \to \Box\varphi)$. Now the relevant φ for $\circ p$ is $\exists q((q \equiv p) \,\&\, F = [\lambda y\, q])$.

Suppose, for reductio, that this particular φ were a rigid condition on properties. Then if we reason as follows (freely using the Rule of Substitution in many of the steps), we can show that our system would yield contradiction:

$\vdash_\Box \forall F(\varphi \to \Box\varphi)$	by definition of strict canonicity
$\vdash \Box\forall F(\varphi \to \Box\varphi)$	by RN
$\vdash \neg\Diamond\neg\forall F(\varphi \to \Box\varphi)$	by Df$\Box$ (158.2)
$\vdash \neg\Diamond\exists F(\varphi \,\&\, \neg\Box\varphi)$	by classical quantifier reasoning
$\vdash \neg\Diamond\exists F(\varphi \,\&\, \Diamond\neg\varphi)$	by (158.11)
$\vdash \neg\exists F\Diamond(\varphi \,\&\, \Diamond\neg\varphi)$	by BF$\Diamond$ (167.3)
$\vdash \forall p\neg\exists F\Diamond(\varphi \,\&\, \Diamond\neg\varphi)$	by GEN
$\vdash \neg\exists p\exists F\Diamond(\varphi \,\&\, \Diamond\neg\varphi)$	by quantifier negation

But the proof of (297) establishes $\vdash \exists p\exists F\Diamond(\varphi \,\&\, \Diamond\neg\varphi)$. So, on pain of system inconsistency, φ fails to be a rigid condition on properties and $\circ p$ fails to be (identical to) a *strictly* canonical object.

For the record, we could have strengthened (297) to the claim:

$$\forall p \exists F \Diamond(\exists q((q \equiv p) \,\&\, F = [\lambda y\, q]) \,\&\, \Diamond\neg\exists q((q \equiv p) \,\&\, F = [\lambda y\, q]))$$

In the proof of (297), we didn't have to choose our witnesses to the existential quantifiers be a contingently true proposition and a propositional property constructed from a necessarily true proposition. We could have chosen the witnesses to be:

- a contingently false proposition and a property constructed from a necessarily false proposition,
- a necessarily true proposition and a property constructed from a contingently true proposition, or
- a necessarily false proposition and a property constructed from a contingently false proposition.

Since every proposition has to be either contingently true, contingently false, necessarily true, or necessarily false, we would have covered all the cases. Thus, we would have a proof of the stronger claim displayed above.

But it suffices, for our purposes to show only (297), since that establishes that the formula φ we're considering isn't a rigid condition on properties. This is instructive because it tells us that any general conclusions we draw about the properties $\circ p$ encodes will rest on a contingency and so such conclusions will fail to be modally strict. The theorems in (299)$\star$ below constitute good examples. If the point comes as a surprise, a closer inspection of the proofs of the theorems that follow should make it clearer. It may also be worth noting that when we define possible worlds in Chapter 12, we shall be in a position to define, for each world w, *the truth-value of p with respect to w*. The truth-value of p with respect to a world w is, by contrast, a strictly canonical object. For more on this world-relativized notion of the truth-value of p, see (557) – (562).

(**299**) $\star$**Lemmas:** Lemmas Concerning Truth-Values of Propositions. The following lemmas are simple consequences of our definitions: (.1) the truth-value of p is a truth-value of p; (.2) the truth-value of p encodes a property F just in case F is identical to *being such that* q, for some proposition q materially equivalent to p; (.3) the truth-value of p encodes proposition r iff r is materially equivalent to p; and (.5) x is a truth-value of p if and only x is identical to the truth-value of p. Formally:

(.1) $\mathit{TruthValueOf}(\circ p, p)$

(.2) $\forall F(\circ pF \equiv \exists q((q \equiv p) \;\&\; F = [\lambda y\, q]))$

(.3) $\circ p\Sigma r \;\equiv\; (r \equiv p)$

(.4) $\mathit{TruthValueOf}(x, p) \;\equiv\; x = \circ p$

These consequences are not modally-strict. **Exercise:** Show that the following variant of (.3), $\circ p\Sigma r \equiv \mathcal{A}(r \equiv p)$, has a modally strict proof.[197]

(**300**) ⋆**Theorem:** The Fregean Biconditional Principle for Truth-Values of Propositions.

$$\circ p = \circ q \;\equiv\; p \equiv q$$

That is, the truth-value of p is identical to the truth-value of q if and only if p is materially equivalent to q.

(**301**) ⋆**Theorem:** The Truth-Value of q is a Truth-Value.

$$\mathit{TruthValue}(\circ q)$$

We next work our way to the modally-strict claim that there are exactly two truth-values.

(**302**) **Definitions and Theorem:** The True and The False. We take advantage of the fact that canonical descriptions are significant (252) to define *The True* ('$\top$') as the abstract object that encodes all and only properties constructed from true propositions, and define *The False* ('$\bot$') as the abstract object that encodes all and only properties constructed from false propositions:

(.1) $\top =_{df} \imath x(A!x \,\&\, \forall F(xF \equiv \exists p(p \,\&\, F = [\lambda y\, p])))$

(.2) $\bot =_{df} \imath x(A!x \,\&\, \forall F(xF \equiv \exists p(\neg p \,\&\, F = [\lambda y\, p])))$

Clearly, both $\top$ and $\bot$ are canonical objects (253), but we leave it as an exercise to show that they are not strictly canonical (260.2).[198] The conditions under which $\top$ and $\bot$ encode properties fail to be rigid and any conclusions we draw about the properties these two objects encode will fail to be modally strict.

[197] Here is a solution. Where φ is the formula $\exists q((q \equiv p) \,\&\, F = [\lambda y\, p])$, then the following is an instance of (258):

$$\imath x(A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])))F \equiv \mathcal{A}\exists q((q \equiv p) \,\&\, F = [\lambda y\, q])$$

So by theorem (296.1), Rule =E, and GEN we know:

(ϑ) $\forall F(\circ pF \equiv \mathcal{A}\exists q((q \equiv p) \,\&\, F = [\lambda y\, q]))$

($\rightarrow$) Assume $\circ p\Sigma r$. Then by definition (295), $\circ p[\lambda y\, r]$. So by (ϑ), $\mathcal{A}\exists q((q \equiv p) \,\&\, [\lambda y\, r] = [\lambda y\, q])$. By (139.10), $\exists q\mathcal{A}((q \equiv p) \,\&\, [\lambda y\, r] = [\lambda y\, q])$. Suppose q_1 is such a proposition, so that we know $\mathcal{A}((q_1 \equiv p) \,\&\, [\lambda y\, r] = [\lambda y\, q_1])$. Then it follows that $\mathcal{A}(q_1 \equiv p)$ and $\mathcal{A}[\lambda y\, r] = [\lambda y\, q_1]$, by (139.2). But the latter implies $[\lambda y\, r] = [\lambda y\, q_1]$, by (175.1). So $r = q_1$, by the definition of proposition identity. Hence $\mathcal{A}(r \equiv p)$.

($\leftarrow$) Assume $\mathcal{A}(r \equiv p)$. Since $[\lambda y\, r] = [\lambda y\, r]$, we know $\mathcal{A}[\lambda y\, r] = [\lambda y\, r]$. So $\mathcal{A}(r \equiv p) \,\&\, \mathcal{A}[\lambda y\, r] = [\lambda y\, r]$. Hence $\mathcal{A}((r \equiv p) \,\&\, [\lambda y\, r] = [\lambda y\, r])$, which implies $\exists q\mathcal{A}((q \equiv p) \,\&\, [\lambda y\, r] = [\lambda y\, q])$. So by ($\vartheta$), $\circ p[\lambda y\, r]$. Hence by (295), $\circ p\Sigma r$. ⋈

[198] In a later chapter, when we introduce possible worlds and use w as a restricted variable to range over possible worlds, we will define The-True-at-w ($\top_w$) and The-False-at-w ($\bot_w$). These distinguished world-relativized truth-values are strictly canonical. See items (559.1) and (559.2).

Nevertheless, one can prove by modally strict means that The True and The False are distinct objects:[199]

(.3) $\top \neq \bot$

(**303**) ⋆**Theorems:** The True and The False Are Truth-Values:

(.1) $TruthValue(\top)$

(.2) $TruthValue(\bot)$

Note that (.1) and (.2) are *not* trivialities. We haven't stipulated that The True and The False are truth-values. Note that one could use (.1) and (.2) to independently prove that there are exactly two truth-values: we would only need to show that any truth-value is identical to $\top$ or $\bot$. But though this would be a perfectly good proof, it wouldn't be a modally strict proof, whereas the proof of (292) is.

(**304**) ⋆**Lemmas:** TruthValueOf, The True, and The False.

(.1) $TruthValueOf(x, p) \rightarrow (p \equiv x = \top)$

(.2) $TruthValueOf(x, p) \rightarrow (\neg p \equiv x = \bot)$

These lemmas need no gloss; they help us to prove the next theorems.

(**305**) ⋆**Theorems:** Facts About $\circ p$, $\top$, and $\bot$. The following two principles governing truth-values are provable: (.1) a proposition is true iff its truth-value is The True; (.2) a proposition is false iff its truth-value is The False:

(.1) $p \equiv (\circ p = \top)$

(.2) $\neg p \equiv (\circ p = \bot)$

These two principles seem obvious *a priori*. But they are capable of proof and so have been proved, in compliance with Dedekind's maxim that serves as the epigraph to Chapter 9.

It is also straightforward to show that: (.3) p is true iff The True encodes p; (.4) p is true iff The False fails to encode p; (.5) $\neg p$ is true iff The False encodes p, and (.6) $\neg p$ is true iff The True fails to encode p, i.e.,

(.3) $p \equiv \top\Sigma p$

(.4) $p \equiv \neg\bot\Sigma p$

[199] An earlier version of this monograph incorrectly annotated the following claim as a ⋆-theorem. I'd like to thank Daniel West for pointing that out and for the detailed sketch he sent, showing how it could be proved by modally strict means. The proof in the Appendix reconstructs the underlying idea and eliminates all the appeals to a Rule of Substitution.

(.5) $\neg p \equiv \bot\Sigma p$

(.6) $\neg p \equiv \neg\top\Sigma p$

Exercises: Show that the following variants of (.3) and (.5), namely $\mathcal{A}p \equiv \top\Sigma p$ and $\mathcal{A}\neg p \equiv \bot\Sigma p$, have modally strict proofs. (Hint: The proofs are analogous to the proof in footnote 197, which contains a solution to the Exercise in (299) of showing that there is a modally strict proof of $\circ p\Sigma r \equiv \mathcal{A}(r \equiv p)$.)

10.2 Extensions of Propositions

Though Frege thought that sentences *denote* truth-values, he never claimed that a truth-value was the *extension* of a sentence. But in Carnap 1947 (26), section numbered 6-1 is titled "The *extension of a sentence* is its truth-value" and section numbered 6-2 is titled "The *intension of a sentence* is the proposition expressed by it". These stipulations on Carnap's part are central to his method of extension and intension (1947, 1–68) and to the semantics of formal logical systems based on that method.

By contrast, we define an *extension of* a proposition p to be an abstract object of a certain kind and show that these abstract objects are truth values. So, object-theoretic extensions apply, in the first instance, to propositions rather than to sentences. In the next section, we investigate the *extension of a property*.

(**306**) **Definitions:** Extension of a Proposition. Let us say that x *is an extension of* p just in case x is an abstract object, x encodes only propositional properties, and x encodes all and only propositions materially equivalent to p:

$$ExtensionOf(x,p) \equiv_{df} A!x \,\&\, \forall F(xF \to Propositional(F)) \,\&\, \forall q((x\Sigma q) \equiv (q \equiv p))$$

No special existence clauses are needed in the definiens.

(**307**) **Theorems:** An Equivalence. It now follows that $ExtensionOf(x,p)$ is equivalent to $TruthValueOf(x,p)$:

$$ExtensionOf(x,p) \equiv TruthValueOf(x,p)$$

Since this equivalence is established by a modally strict proof, a Rule of Substitution allows us to substitute one for the other wherever either occurs as a subformula. Note also that if we apply GEN and the Rule of Actualization to this theorem, then it follows by (149.1) that, for any individual x, x is identical to $\imath xExtensionOf(x,p)$ iff x is identical to $\imath xTruthValueOf(x,p)$.

(**308**) **Theorems:** Fundamental Theorems of Extensions of Propositions. It is now provable that: (.1) there is a unique extension of p; (.2) the extension of p exists; (.3) the extension of proposition p is the truth-value of p, and (.4) the extension of proposition p is a truth-value:

(.1) $\exists!x ExtensionOf(x,p)$

(.2) $\imath x ExtensionOf(x,p)\downarrow$

(.3) $\imath x ExtensionOf(x,p) = \circ p$

Cf. (.3) with Carnap's assertion (1947, 26) that the extension of a sentence is its truth-value. On our reconstruction, this is a principle is about propositions, not sentences.

10.3 Extensions of Properties: Natural Classes

(**309**) **Remark:** Natural vs. Theoretical Mathematics. In what follows it is important to distinguish between *natural* mathematics and *theoretical* mathematics. Natural mathematics consists of ordinary, pretheoretic claims we make about mathematical objects, such as the following:

- The *number* of planets is eight.
- There are more individuals in the *class* of insects than in the *class* of humans.
- The lines on the pavement have the same *direction*.
- The figures drawn on the board have the same *shape*.

By contrast, the claims of theoretical mathematics are the axioms, theorems, hypotheses, conjectures, etc., asserted in the context of some explicit mathematical theory or in the context of some implicit or informal, but distinctly mathematical, assumptions. Example of such claims are:

- In Zermelo-Fraenkel set theory, the null set is an element of the unit set of the null set.
- In Real Number Theory, 2 is less than or equal to π.

The sentence operator "In theory T" is frequently omitted when mathematicians make such claims.

One distinguishing feature of pure theoretical mathematics is that the fundamental axioms and assumptions of those theories govern special, abstract mathematical relations and operations (e.g., membership, predecessor, less than, addition, etc.) and don't involve ordinary relations or individuals.[200] An analysis of theoretical mathematics is reserved for Chapter 15, where its

[200] Of course, there may be non-mathematical relations and assumptions in impure or applied mathematics, but such relations won't be subject to our analysis; rather they are subject to the investigations of natural science.

objects and relations are identified as abstract objects and abstract relations, respectively. Prior to that chapter, however, we shall have occasion to analyze various *natural* mathematical objects, by identifying them as abstractions from the body of ordinary (exemplification) predications that are independent of any mathematical theory.

The first group of natural mathematical objects we examine are the natural classes. A natural class is not an abstraction based on the axioms of some mathematical theory of sets, but is rather an abstraction from the facts about the exemplification of properties (both as to how they are exemplified by different individuals and how different properties may be exemplified by the very same individuals). In what follows, then, we shall be investigating *natural* classes and analyzing them as the *extensions* of properties.[201]

(**310**) **Remark:** In Sections 10.1 and 10.2, we introduced the notion of a *truth-value* and then the notion an *extension of a proposition*. In this section, however, we first define the notion of an (exemplification) *extension of a property* and subsequently use it to define the notion a (natural) *class*. The change in the order of presentation can be understood as follows. Frege (Frege 1891) introduced the notion of a *truth-value* prior to Carnap's claim (1947, 26) that the extension of a sentence is its truth value. But Carnap extended the historical notion of an *extension*, which had traditionally been applied to general terms (see below), to sentences. So we introduced truth values, then extensions of propositions, and then validated a version of Carnap's insight in the present system.

But the notion of an *extension of a general term* goes back to medieval logic and is well-entrenched in the Port Royal Logic (1662), whereas the notion of a *class* is much more modern.[202] We therefore define the modern notion of a *class* in terms of the older philosophical notion, though in our system, we take *extensions* to apply, in the first instance, to properties rather than to general terms.

Our analysis retains some features of Whitehead & Russell's *no class* theory (1910–1913 [1925–1927], $*20{\cdot}02$ and $*20{\cdot}3$), in so far as it attempts to introduce classes by definition. We'll also prove a theorem (368)★ that bears a similarity to their theorem $*20 \cdot 3$ (and corresponding definition $*20 \cdot 01$). But, unlike the *no class* theory, we explicitly define classes rather than contextually define

[201] If we think semantically for the moment and allow ourselves some primitive set theory, then properties would naturally be assigned *two* extensions – an exemplification extension (which could vary from world to world) and an encoding extension. The exemplification extension of F is the set of objects x that make Fx true and the encoding extension of F is the set of objects that make xF true. The natural classes we now plan to investigate are the objects that represent the *exemplification* extensions of properties.

[202] See Buroker 2014 (Section 3), for a discussion of the notion of an *extension of a general term* in the Port Royal Logic.

them. Before we develop our analysis, we begin with an important distinction.

(**311**) **Remark:** The Naive/Logical Conception of Set or Class. The conception of natural classes just described is very closely related to the *naive conception of set*, which derives from Cantor. In the opening lines of Cantor 1895, 481 (1915, 85; 1932, 282), we find:[203]

> By a 'set' we understand any collection into a whole M of definite well-differentiated objects m of our intuition or thought (which are called the 'elements' of M).

Moreover, in Cantor 1932, 204, we find:[204]

> By a 'manifold' or 'set' I understand a many which can be thought of as one, i.e., a totality of particular elements that can be combined into a whole by a law, and I believe something is defined thereby that is related to Platonic Forms or Ideas, ...

The final clause of the quoted passage, in which Cantor relates sets to Platonic Forms, is interesting. We'll see, in Chapter 11, that for each property F, the Form of F (Φ_F) can be defined as abstract object. But one way to interpret Cantor here is that that for each property F, the 'totality' of which Cantor speaks is the *extension* of F and the 'law that combines into a whole' is the axiom or definition that states the exemplification conditions of F (i.e., the principle that provides the exemplification conditions for something to be in the extension of F). Then our work in Chapter 11 will indeed show that sets bear a connection to Platonic Forms.

Boolos elaborates on the naive conception of set as follows (1971, 216):[205]

[203] The original is:

> Unter einer 'Menge' verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die 'Elemente' von M genannt werden) zu einem Ganzen.

[204] The original is:

> Unter einer "Mannigfaltigkeit" oder "Menge" verstehe ich nämlich allgemein jedes Viele, welches sich als Eines denken läßt, d.h. jeden Inbegriff bestimmter Elemente, welcher durch ein Gesetz zu einem Ganzen verbunden werden kann, und ich glaube hiermit etwas zu definieren, was verwandt ist mit dem *Platonischen* εἶδος oder ἰδέα, ...

[205] Later, Boolos contrasts the naive conception with the *iterative* conception, where sets are structured in such a way that they arise in an ordered series of stages. In describing the iterative conception, Boolos writes (1971, 221):

> At stage zero, there is a set for each possible collection of individuals (and if there are no individuals, there is only one set, namely, the null set). At stage one, there exists a set for each possible collection consisting of individuals and sets formed at

> Here is an idea about sets that might occur to us quite naturally, and is perhaps suggested by Cantor's definition of a set as a totality of definite elements that can be combined into a whole *by a law*.
>
> By the law of excluded middle, any (one-ary) predicate in any language either applies to a given object or does not. So, it would seem, to any predicate there correspond two sorts of thing: the sort of thing to which the predicate applies (of which it is true) and the sort of thing to which it does not apply. So, it would seem, for any predicate there is a set of all and only those things to which it applies (as well as a set of just those things to which it does not apply). Any set whose members are exactly the things to which the predicate applies—by the axiom of extensionality, there cannot be two such sets—is called the *extension* of the predicate. Our thought might therefore be put: "Any predicate has an extension." We shall call this proposition, together with the argument for it, the *naive conception of set*.

Boolos then generalizes further by moving from property terms to open formulas φ in which y doesn't occur free (but which typically have a free occurrence of x). Using $\mathcal{K}$ to denote a standard first-order language having (a) variables that range over both sets and individuals, (b) a distinguished property term S for *being a set*, and (c) a distinguished binary relation term $\in$ for *membership*, Boolos writes (1971, 217):

> If the naive conception of set is correct, there should (at least) be a set of just those things to which φ applies, if φ is a formula of $\mathcal{K}$. So (the universal closure of) $\ulcorner(\exists y)(Sy \,\&\, (x)(x \in y \equiv \varphi))\urcorner$ should express a truth about sets (if no occurrence of 'y' in φ is free).

Of course, Boolos takes it that he has properly represented the central claim ("every predicate has an extension") of the naive conception because he is assuming that every open formula φ with free variable x defines a property, either by way of the λ-expression $[\lambda x\, \varphi]$ or by way of the instance $\exists F \forall x(Fx \equiv \varphi)$ of property comprehension.[206]

But this is an assumption that object theory doesn't completely *endorse*. In object theory, though every open formula φ with a free occurrence of an individual variable y can be used to formulate a λ-expression $[\lambda y\, \varphi]$, not every such expression is guaranteed to be significant. Axiom (39.2), for example,

stage zero. And so on, until one reaches stage omega, at which there exists a set for each possible collection consisting of individuals and sets formed at stages one, two, three, ... Of course, this is only the beginning.

We shall put aside further discussion of the iterative conception of set, since that conception informs our understanding of the theoretical mathematics of sets and, in particular, Zermelo-Fraenkel set theory (ZF). The philosophical analysis of the language and theorems of ZF will be discussed in Chapter 15, where we analyze theoretical mathematics generally.

[206] See also Cocchiarella 1986, 1988, who also distinguishes a 'logical' notion of a class.

only guarantees that $[\lambda x\, \varphi]$ is significant if it is a core λ-expression. Axiom (49) guarantees that $[\lambda x\, \varphi]$ is significant if for some formula ψ, both $[\lambda x\, \psi]$ is significant and $\Box\forall x(\psi \equiv \varphi)$. And the Kirchner Theorem (271.1) guarantees that a λ-expression is significant if and only if, necessarily, its matrix can't distinguish indiscernible objects. These principles forestall the paradoxes of naive object theory. And we'll see that they additionally forestall the paradoxes of naive set theory once we formally define the logical notion of a set and define *member of*. We shall nevertheless be able to approximate, in the material mode, the central claim of the naive conception of set ("every predicate has an extension") as a theorem. We shall do this by defining the notion of $Class(x)$ and then deriving: for every property F, there is a class whose elements are precisely the individuals exemplifying F, i.e.,

> **Fundamental Principle of Naive Set Theory**
> $\forall F \exists x(Class(x) \,\&\, \forall y(y \in x \equiv Fy))$

See (318) below. Thus, Boolos' talk of predicates and predicate application becomes represented by talk of properties and property exemplification, and so the claim "every predicate has an extension" becomes represented as the principle: for every property F, there is a class whose members are precisely the individuals exemplifying F. Indeed, we'll see in Remark (313) below that one may even preserve the idea that every 'predicate' has an extension, if one accepts that λ-expressions (i.e., complex unary relation terms) that aren't significant have an empty extension.

If this is correct, then the natural conception of a class and the logical conception of set nicely dovetail and can be formalized together. One of our goals in what follows, therefore, is to show that this analysis is indeed correct. To do this, we: (a) precisely define what it is for an abstract object to be an extension of a property, (b) define classes to be extensions of properties, (c) prove that *every* property has an extension, (d) prove that every class is the extension of some property, (e) define membership in a class, and (f) prove that for every property F, there is a class whose members are precisely the individuals exemplifying F. Since these and other principles formulated below make the naive conception of set formally precise, we shall henceforth re-label the 'naive' conception with the following, less rhetorical label: the *logical conception* of set. Thus, in what follows, natural classes are identified with sets logically conceived.

10.3.1 Basic Definitions and Theorems

(**312**) **Definition:** Extension Of, Class Of, and Class. We now say that x is an (exemplification) *extension of* G, or x is a *class of* Gs, if and only if x is an abstract object, G exists, and x encodes just the properties materially equivalent to G (with respect to exemplification):

$$(.1)\quad \left.\begin{array}{c} \mathit{ExtensionOf}(x,G) \\ \mathit{ClassOf}(x,G) \end{array}\right\} \equiv_{df} A!x \;\&\; G\downarrow \;\&\; \forall F(xF \equiv \forall z(Fz \equiv Gz))$$

In the usual way, this is shorthand for a definition that uses metavariables κ and Π^1 instead of the object language variables x and G.[207]

Whereas Carnap famously stipulated that the extension of a predicate is a class (1947, 19, 4-14), we adopt the view that a natural class may be defined as an extension of a property. Accordingly, we say (.2) x is a (natural) *class*, or x is a *logical set*, if and only if x is an extension of (a or class of) some property G:

$$(.2)\quad \left.\begin{array}{c} \mathit{LogicalSet}(x) \\ \mathit{Class}(x) \end{array}\right\} \equiv_{df} \left\{\begin{array}{c} \exists G(\mathit{ExtensionOf}(x,G)) \\ \exists G(\mathit{ClassOf}(x,G)) \end{array}\right.$$

For simplicity, we've used $\mathit{Class}(x)$ instead of $\mathit{NaturalClass}(x)$ as one of the definienda. We don't need $x\downarrow$ in the definiens, by the following reasoning. If it is known, by proof or by hypothesis, that κ is a non-denoting individual term, then $\neg\mathit{ExtensionOf}(\kappa,G)$ becomes derivable, for every G. Hence we can derive $\neg\exists G(\mathit{ExtensionOf}(x,G))$, and so $\neg\mathit{Class}(\kappa)$ when κ is any term known to be empty. The clause $x\downarrow$ is therefore not needed in the definiens.

(**313**) **Remark:** Intuitions About What Extensions Are. In the definiens of the foregoing definitions, none of the notions of collection, whole, totality, element, membership, etc., are used. Nevetheless, the definition isn't completely unrelated to Cantor's conception of a set as a collection of definite entities combined into a whole by a law (quoted above). Let us analyze Cantor's conception as involving a *law*, *a collection of entities*, and the *combining into a whole*. On the above definition, a class is an extension of some property G, where an extension is an abstract object that encodes the properties F such that $\forall F(Fz \equiv Gz)$. So, for any property G, the formula $\forall z(Fz \equiv Gz)$ constitutes a rule, or *law*-like condition on properties, that defines a *collection of properties F* which comply with the law. Any x that encodes all and only such properties F, is an object that *combines into a whole* the collection of those properties F that satisfy the law. Thus, whereas Cantor conceived of the set of Gs *from below*, as the totality of objects x obeying the rule Gx, our more Fregean conception carves out

[207] Since we can use bound object-language variables in the definiens (given that alphabetically-variant formulas are inter-derivable), the definition just formulated in the text is short for:

$$\left.\begin{array}{c} \mathit{ExtensionOf}(\kappa,\Pi^1) \\ \mathit{ClassOf}(\kappa,\Pi^1) \end{array}\right\} \equiv_{df} A!\kappa \;\&\; \Pi^1\downarrow \;\&\; \forall F(\kappa F \equiv \forall z(Fz \equiv \Pi^1 z))$$

provided F and z don't occur free in Π^1. This ensures (a) that there is an instance of the definition for every individual term κ and unary relation term Π, and (b) that the definiens, and hence the definiendum, will be false whenever either κ or Π^1 is an empty term.

the extension of G *from above*, as the reified totality of properties that obey the higher-order rule: being a property exemplified by the same objects that exemplify G.[208] In what follows, we'll *define* a member or element of the extension of G to be any individual x such that Gx. But the extension itself encodes properties, in the first instance, and derivatively contains individuals as members.

If we temporarily allow ourselves some intuitive notions from set theory, then one could say that, for each G, the class of Gs objectifies the equivalence class of properties F that are materially equivalent to G. So the x in the definition of $ExtensionOf(x,G)$ is an abstraction over the properties that are in the cell of the partition that contains G.

Another observation about extensions concerns the temptation to eliminate the conjunct $G\downarrow$ from definition (312), so that the definition, in the case of extensions, becomes:

$$(\vartheta)\ \ ExtensionOf(x,G) \equiv_{df}\ A!x \,\&\, \forall F(xF \equiv \forall z(zF \equiv zG))$$

Given our conventions for definitions, we could instantiate (ϑ) to property terms that aren't significant. Then, where Π is such a term (because, say, $\vdash \neg\Pi\downarrow$), it would follow that the abstract object that encodes all and only properties F that are unexemplified is the extension of Π. For those are precisely the properties that would satisfy the formula $\forall z(Fz \equiv z\Pi)$. And, when we define membership in (316) below, (ϑ) would yield that the extension of Π has no members. Thus, one might reformulate definition (312) so that property terms that aren't significant have an extension.

But this option will not be pursued in what follows, for several reasons. First, we discussed in Remark (36) how the inclusion of claims like $G\downarrow$ in the definitions-by-$\equiv$ forestalls violations of the *garbage in*, *garbage out* principle. We prefer to avoid the situation in which $ExtensionOf(x,\Pi)$ is true when Π doesn't have a denotation. Moreover, (ϑ) would lead to the introduction of an *impractical term*, as discussed in Remark (283). When we introduce the operator $\epsilon(\,)$ in (322) below, (ϑ) would yield terms the form $\epsilon\Pi$ that are significant even though the argument Π is not. Though the discussion in Remark (283) explains why these don't pose a logical problem, we can avoid introducing such impractical terms in this case by including the conjunct $G\downarrow$ in definition (312).

One final note concerns the question of whether we should also define an

[208] Note that it is *not* an option to define an extension of G as encoding all and only those Fs that are the haecceities of objects that exemplify G. That is, we may not define:

$$(\xi)\ \ ExtensionOf(x,G) =_{df} A!x \,\&\, G\downarrow \,\&\, \forall F(xF \equiv \exists y(Gy \,\&\, F = [\lambda z\, z = y]))$$

We know, from the McMichael-Boolos paradox, that $[\lambda z\, z = y]$ doesn't signify a property when y is an abstract object. So if G is exemplified by some abstract object y, $F = [\lambda z\, z = y]$ is false for every F. Hence, given (ξ), the null object would be the extension of such a G, contrary to intuition. So we won't be using that definition.

encoding extension of G to be an abstract object that encodes exactly the properties F that are encoded by the same objects as G, as in:

$$EncodingExtensionOf(x,G) \equiv_{df} A!x \,\&\, G\downarrow \,\&\, \forall F(xF \equiv \forall z(zF \equiv zG))$$

But since $\forall z(zF \equiv zG) \equiv F = G$ (189), an encoding extension of G would encode only G and no other properties. Intuitively, this definition would partition the domain of properties into equivalence classes each of which contains a single property. We won't be exploring this development in what follows, though see the discussion of the *thin Form* of G in the next chapter.

(**314**) **Theorem:** Pre-Basic Law V and Other Facts. It is a modally strict fact underlying Frege's Basic Law V that (.1) if x is an extension of G and y is an extension of H, then x is identical to y if and only if G and H are exemplified by the same objects:

(.1) $(ExtensionOf(x,G) \,\&\, ExtensionOf(y,H)) \rightarrow (x = y \equiv \forall z(Gz \equiv Hz))$

It is also a useful fact that (.2) if x is an extension of some property H and encodes both F and G, then F and G are materially equivalent:

(.2) $(ExtensionOf(x,H) \,\&\, xF \,\&\, xG) \rightarrow \forall z(Fz \equiv Gz)$

It follows straightforwardly that (.3) if an object x encodes properties that aren't materially equivalent, then x is not a class:

(.3) $(xF \,\&\, xG \,\&\, \neg\forall z(Fz \equiv Gz)) \rightarrow \neg Class(x)$

(**315**) **Theorems:** There is a (Unique) Extension/Class of G and the Existence of Classes.

(.1) $\begin{cases} \exists x ExtensionOf(x,G) \\ \quad \exists x ClassOf(x,G) \end{cases}$

(.2) $\begin{cases} \exists! x ExtensionOf(x,G) \\ \quad \exists! x ClassOf(x,G) \end{cases}$

It follows from (.2) by GEN that for every property G, there is a unique extension of G. This consequence doesn't quite yet capture the intuition that every property has an extension whose members are precisely the objects exemplifying the property, since we haven't yet defined membership or shown that if x is an extension of G, then the members of x are all and only the objects that exemplify G. However, from (.2) we can derive that (.3) the extension of G exists:

(.3) $\iota x ExtensionOf(x,G)\downarrow$

In the next section, we'll use this last fact to reconstruct a Fregean operator that takes a property term Π as an argument and yields a term that denotes the extension of the property denoted by Π if Π denotes a property, and denotes nothing otherwise.

It is a straightforward consequence of (.1) and definition (312.2) that (.4) classes exist:

(.4) $\exists x Class(x)$

(**316**) **Definition:** Membership. We now say that y is a *member of* x, or y is an *element of* x, written $y \in x$, if and only if x is an extension of some property that y exemplifies:

$$y \in x \ \equiv_{df} \begin{cases} \exists G(ExtensionOf(x,G) \,\&\, Gy) \\ \quad \exists G(ClassOf(x,G) \,\&\, Gy) \end{cases}$$

This definition of membership is similar to, but doesn't quite match, Frege's. Frege's definition (1893, §34, A) is discussed in some detail below, in Remark (332), and we'll see there that the above differs from Frege's definition in a number of ways.

(**317**) **Theorem:** Membership and Exemplification. A precise correlation between membership and exemplification now obtains for extensions, namely, (.1) if x is an extension of H, then y is an element of x iff y exemplifies H

(.1) $ExtensionOf(x,H) \to \forall y(y \in x \equiv Hy)$

From this, it is straightforward to show that (.2) the Russell property, being an x that is a member of itself, doesn't exist:

(.2) $\neg[\lambda x\, x \notin x]\downarrow$

Hence, we may not instantiate $[\lambda x\, x \notin x]$ into the universal generalization of theorem (315.1) to obtain that $\exists x ExtensionOf(x,[\lambda x\, x \notin x])$. Indeed, it is provable that:

(.3) $\neg\exists x ExtensionOf(x,[\lambda x\, x \notin x])$

(**318**) **Theorem:** The Fundamental Theorem for Natural Classes and Logically-Conceived Sets. We saw in Remark (311) that the fundamental principle governing the conception of natural classes and logical sets is: for every property F, there is a natural class (i.e., logical set) whose members are precisely the individuals exemplifying F. We now have a proof of this claim:

$$\forall F \exists x(Class(x) \,\&\, \forall y(y \in x \equiv Fy))$$

The proof in the Appendix shows how this follows from previous theorems.

(**319**) **Theorem:** Existence of a Self-Membered Natural Class.[209] It is a significant fact that there exists a natural class that is a member of itself:

$\exists x(Class(x) \,\&\, x \in x)$

10.3.2 Natural Classes, Logical Sets, and Modality

(**320**) **Theorem:** Natural Classes Aren't Necessarily Classes. Recall that in (221.1) we proved that there are properties that are materially equivalent but possibly not materially equivalent, and that in (221.3) we proved that there are properties that are actually materially equivalent but possibly not materially equivalent. In light of this, let's extend our earlier observation that appealed to some intuitive set theory, on which the class of Gs objectifies the equivalence class of properties materially equivalent to F. If we now appeal to the intuitive (i.e., not yet defined) notion of possible world, then we could say: since the equivalence classes of materially equivalent properties vary from world to world, the abstractions that arise from materially equivalent properties at one world are not the same as those that arise from materially equivalent properties at another world. Indeed, given any property F whose exemplification extension varies from world to world, the object abstracted from the properties materially equivalent to F at one world won't be the same object as the ones abstracted from the properties materially equivalent to F at other worlds – they will encode different properties.

These observations may help one to appreciate the following interesting facts. The first is that (.1) it is not the case that for every x and G, if x is an extension of G, then necessarily x is an extension of G:

(.1) $\neg\forall x\forall G(ExtensionOf(x,G) \rightarrow \Box ExtensionOf(x,G))$

The second interesting fact is (.2) not every (natural) class is necessarily a (natural) class:

(.2) $\neg\forall x(Class(x) \rightarrow \Box Class(x))$

Thus, some classes are not necessarily classes. The final interesting fact is (.3) it is not necessary that every class is actually a class:

(.3) $\neg\Box\forall x(Class(x) \rightarrow \mathcal{A}Class(x))$

(.2) and (.3) have some interesting consequences when we start reasoning with restricted variables, especially in modal contexts. This will be discussed in Section 10.5, in (341).

[209] I'd like to thank Daniel West for suggesting that I add this theorem.

Recall also that in (221.4) we proved that for every property F, there is a property materially equivalent to F but possibly not. From this, we can prove a claim even stronger than (.1) – (.3), namely, that (.4) if x is an extension of H, then x might not be an extension of H, and (.5) classes (in general) possibly fail to be classes:

(.4) $\mathit{ExtensionOf}(x,H) \rightarrow \Diamond\neg\mathit{ExtensionOf}(x,H)$

(.5) $\mathit{Class}(x) \rightarrow \Diamond\neg\mathit{Class}(x)$

It follows trivially from (.5) that natural classes aren't necessarily classes! This is a fact about *natural* classes abstracted, in part, from contingent exemplification predications. For an alternative conception of classes, one defined by axioms governing the mathematically *primitive* notion of set membership, we have to turn to *theoretical* mathematics and (axiomatic) set theory. This takes placve in Chapter 15

(**321**) **Theorem:** Membership is not a Necessary Condition. It is important to observe that the condition $y \in x$, as defined in (316), does *not* generally hold by necessity when it holds – it is not the case that for every x and y, if x is a member of y then necessarily x is a member of y:

(.1) $\neg\forall x\forall y(y \in x \rightarrow \Box y \in x)$

The proof of this claim in the Appendix appeals only to resources within the system. But if we temporarily (a) *extend* our system with some plausible (modal) facts, such as that Socrates is wise (Ws) but might not have been ($\Diamond\neg Ws$), and (b) appeal to the semantically primitive notion of a possible world, we can give an intuitive argument for this theorem by showing why it holds in models where the modal facts hold. In such models, we can find an x and y such that $y \in x$ holds at the actual world $\boldsymbol{w}_0$, but fails to hold at some other possible world. It is to be emphasized, however, that the proof of this theorem given in the Appendix makes no use of the modal facts in (a), but rather analogous modal facts available within the system.

Here then is the semantic proof sketch. By hypothesis, Ws is true at $\boldsymbol{w}_0$ but false at some other possible world, say $\boldsymbol{w}_1$. Now consider any x such that, at $\boldsymbol{w}_0$, $\mathit{ExtensionOf}(x,W)$. Then we know several things: by definition of $\in$, it follows that (a) $s \in x$ holds at $\boldsymbol{w}_0$, and by definition of $\mathit{ExtensionOf}$, it follows that (b) x encodes at $\boldsymbol{w}_0$ all and only the properties materially equivalent to W at $\boldsymbol{w}_0$. Hence, it follows from (b) that x encodes W at $\boldsymbol{w}_0$ and since encoding is rigid, it follows that (c) x encodes W at $\boldsymbol{w}_1$.

Now to see that $\neg(s \in x)$ holds at $\boldsymbol{w}_1$, suppose, for reductio, that $s \in x$ holds at w_1. Then, $\exists G(\mathit{ExtensionOf}(x,G)\ \&\ Gs)$ holds at $\boldsymbol{w}_1$, by definition. Let P be such a property, so that we know both $\mathit{ExtensionOf}(x,P)$ and Ps hold at $\boldsymbol{w}_1$.

The former implies (d) x encodes at $\boldsymbol{w}_1$ all and only the properties materially equivalent to P at $\boldsymbol{w}_1$. But, by hypothesis, $\neg Ws$ at $\boldsymbol{w}_1$, and so W and P are not materially equivalent at $\boldsymbol{w}_1$. Hence, by (d), x doesn't encode W at $\boldsymbol{w}_1$. But by (c), x encodes W at $\boldsymbol{w}_1$. Contradiction.

Indeed, as one might expect, an even stronger claim can be established, namely, that (it is generally the case that) membership, when it holds, fails to hold necessarily:

(.2) $y \in x \rightarrow \neg\Box y \in x$

10.4 Reconstructing *the* Fregean Extension of *F*

(**322**) **Definition:** Notation for the Extension of a Property. Recall that for an arbitrary property G, we established that the description $\iota x ExtensionOf(x,G)$ is significant, by a modally strict proof (315.3). So let's introduce ϵG to rigidly refer to *the extension of* G:

$$\epsilon G \;=_{df}\; \iota x ExtensionOf(x,G)$$

Note that whereas Frege's $\dot{\varepsilon}$-operator is a *variable-binding* operator, our ϵ-operator is not; it is simply a functional, term-forming operator; it can operate on any property term Π. We'll say more about how our operator compares with Frege's in (330) and (332) below.

(**323**) **Theorems:** The ϵ-Operator Doesn't Yield Impractical Terms. It is important to observe that if Π is an empty property term, then $\iota x ExtensionOf(x,\Pi)$ is an empty term, and so is the defined term $\epsilon\Pi$:

(.1) $\neg\Pi\downarrow \rightarrow \neg\iota x ExtensionOf(x,\Pi)\downarrow$, where Π is any unary relation term in which x doesn't occur free

(.2) $\neg\Pi\downarrow \rightarrow \neg\epsilon\Pi\downarrow$, where Π is any unary relation term

Thus, the ϵ-operator doesn't give rise to any *impractical* terms, i.e., complex terms that denote even though the argument to the *defined* term-forming operator is empty. As discussed in the latter part of Remark (283), though our system doesn't rule out such impractical terms, they don't occur in this case because $ExtensionOf(x,\Pi)$ fails to be true when Π is empty. Consequently, the derived Rule of Identity by Definition (120.1) allows us to immediately conclude that $\epsilon\Pi = \iota x ExtensionOf(x,\Pi)$ when Π is significant, and the primitive Rule of Definition by Identity (73.1) yields $\neg\epsilon\Pi\downarrow$ when it is not.[210]

[210] Note that if we had defined $ExtensionOf(x,G)$ more simply as $A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz))$, then there would have been property terms Π such that $\neg\Pi\downarrow \,\&\, \epsilon\Pi\downarrow$. Under this simpler definition of $ExtensionOf(x,G)$, $\epsilon\Pi$ would have denoted (when Π is empty) the abstract object that encodes all and only the unexemplified properties.

(**324**) **Theorem:** The Extension of G is Canonical. It now follows, for an arbitrary property G, that (.1) the extension of G is identical to the abstract object that encodes exactly the properties that are materially equivalent to G:

(.1) $\epsilon G = \imath x(A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz)))$

So by (253), ϵG is (identical to) a canonical individual. Moreover, it can be shown by modally strict means that (.2) the extension of G encodes G:[211]

(.2) $\epsilon G G$

Note the convention of putting the first occurrence of 'G' in a slightly smaller font, to make it easier to parse such formulas as encoding formulas of the form xG with the complex individual term ϵG substituted for x.

(**325**) **Theorem:** Another Fact About Contingently Equivalent Properties. Recall again that in (221.1) we proved that there are properties that are materially equivalent but possibly not materially equivalent. It is straightforward consequence of this that there are properties F and G such that possibly, F and G are materially equivalent but possibly not materially equivalent:

$$\exists G \exists F \Diamond(\forall z(Fz \equiv Gz) \,\&\, \Diamond\neg\forall z(Fz \equiv Gz))$$

This theorem will help us show that ϵG is not a strictly canonical object. Where φ is the formula $\forall z(Fz \equiv Gz)$, the above establishes that there are properties G and F such that $\Diamond(\varphi \,\&\, \Diamond\neg\varphi)$, i.e., by (165.11), that $\Diamond\varphi \,\&\, \Diamond\neg\varphi$.

(**326**) **Remark:** ϵG is Not Strictly Canonical. By (324.1), we know ϵG is (identical to) a canonical object. But it isn't too hard to see that ϵG isn't (identical to) a strictly canonical object. We have to show that where φ is the formula $\forall z(Fz \equiv Gz)$, that φ isn't a rigid condition on properties, i.e., by (260.2), that there is no modally strict proof of $\forall F(\varphi \to \Box\varphi)$. By reasoning analogous to that displayed in Remark (298), we can argue that the assumption that φ is rigid would require our system to be inconsistent:

$\vdash_\Box \forall F(\varphi \to \Box\varphi)$	assumption, definition of strict canonicity
$\vdash \Box\forall F(\varphi \to \Box\varphi)$	by RN
$\vdash \neg\Diamond\neg\forall F(\varphi \to \Box\varphi)$	by Df$\Box$ (158.2)
$\vdash \neg\Diamond\exists F(\varphi \,\&\, \neg\Box\varphi)$	by classical quantifier reasoning
$\vdash \neg\Diamond\exists F(\varphi \,\&\, \Diamond\neg\varphi)$	by (158.11)
$\vdash \neg\exists F\Diamond(\varphi \,\&\, \Diamond\neg\varphi)$	by BF$\Diamond$ (167.3)
$\vdash \forall G\neg\exists F\Diamond(\varphi \,\&\, \Diamond\neg\varphi)$	by GEN
$\vdash \neg\exists G\exists F\Diamond(\varphi \,\&\, \Diamond\neg\varphi)$	by quantifier negation

[211] Thanks to Daniel West for noting, with respect to an earlier draft of this monograph, that this theorem was incorrectly tagged as non-modally strict.

But the proof of (325) establishes $\vdash \exists G \exists F \Diamond(\varphi \,\&\, \Diamond\neg\varphi)$. So, on pain of system inconsistency, φ fails to be a rigid condition on properties and ϵG fails to be (identical to) a *strictly* canonical object.

(**327**) ⋆**Lemmas:** Facts About the Extension of a Property. The previous definition, together with earlier theorems, straightforwardly yield, as non-modally strict theorems, that: (.1) the extension of G is an extension of G; (.2) the extension of G encodes a property F iff F is materially equivalent to G; and (.3) x is an extension of G iff x is identical to the extension of G:

(.1) $ExtensionOf(\epsilon G, G)$

(.2) $\forall F(\epsilon G F \equiv \forall z(Fz \equiv Gz))$

(.3) $ExtensionOf(x, G) \equiv x = \epsilon G$

We've again used the convention using a slightly smaller font for the symbol 'G' on occasion, to make it easier to parse certain formulas.

(**328**) ⋆**Theorem:** Frege's Basic Law V. A consistent version of Frege's Basic Law V is now derivable, namely, the extension of F is identical with the extension of G iff all and only the objects that exemplify F exemplify G:

$$\epsilon F = \epsilon G \equiv \forall x(Fx \equiv Gx)$$

Cf. Frege, *Grundgesetze I*, §20. See Remark (335) for a discussion about how this ⋆-theorem can be turned into a modally strict one. Also, it should be mentioned that in a later chapter, we'll develop a modally strict proof of: the extension of F at w is identical to the extension of G at w if and only if F and G are materially equivalent at w. See (567).

(**329**) **Remark:** Fregean *Epsilon* Notation. In what follows, the reader should carefully distinguish the symbols $\in$, ϵ, and ε. We've introduced the symbol $\in$ for the defined notion of *membership*, as in the formula $x \in y$ (316). And we've introduced ϵ as a term-forming operator, as in ϵF (322). We shall now start using ε to represent Frege's variable-binding *epsilon* operator, in expressions such as $\dot{\varepsilon} f(\varepsilon)$. Since the use of the modern, stylized $\in$ symbol for set membership is historically connected to Peano's use of a Greek *epsilon* (Peano 1889, vi, x; Russell 1903, 19), we are therefore using this Greek letter in at least three different font faces. Fortunately, it is easy to tell them apart.

(**330**) **Remark:** Frege on the Contradiction. After receiving the letter from Russell (dated 16 June 1902) that outlined the paradox that affects predication and set membership, Frege added an Appendix to his (then forthcoming) 1903a and reconstructed the paradox in two different ways: the first derivation doesn't use the defined notion of *membership* whereas the second one does. So let's focus on the first derivation; later, in (332), we'll compare his notion of membership with ours.

In 1903a ([2013, 256]), Frege explicitly notes that the following formula says that Δ is a class that does not belong to itself:

$$\neg\overset{\mathfrak{g}}{\smile}\begin{array}{l}\mathfrak{g}(\Delta)\\ \dot{\varepsilon}(-\!\!-\ \mathfrak{g}(\varepsilon)) = \Delta\end{array}$$

But strictly speaking, this formula says: not every concept $\mathfrak{g}$ is such that its extension Δ falls under it, i.e., some concept $\mathfrak{g}$ has an extension (Δ) that doesn't fall under $\mathfrak{g}$, i.e., Δ is the extension of a concept under which Δ does not fall. Frege then goes on to show how a contradiction can be derived using this formula, by building a paradoxical extension from the concept it expresses.

Frege's formula can be reconstructed in our system, but it doesn't lead to a contradiction because it can't be used to construct an expression for a paradoxical property. If we ignore all the judgment strokes and write his formula using our notation for negations, conditionals, quantifiers, and identity, we have:

$$\neg\forall\mathfrak{g}(\dot{\varepsilon}\mathfrak{g}(\varepsilon)=\Delta \rightarrow \mathfrak{g}(\Delta))$$

Now if we take the liberty of representing Fregean concepts as properties (Frege 1892, 51), use upper case G for the property variable $\mathfrak{g}$, and use our notation ϵG instead of $\dot{\varepsilon}\mathfrak{g}(\varepsilon)$ to refer to the extension of G, the above formula converts to:

$$\neg\forall G(\epsilon G=\Delta \rightarrow G(\Delta))$$

If we use x instead of Δ and exemplification instead of functional application, this becomes:

$$\neg\forall G(\epsilon G=x \rightarrow Gx)$$

But by quantifier laws and the symmetry of identity, this is equivalent to:

$$\exists G(x=\epsilon G \,\&\, \neg Gx)$$

This formula explicitly asserts: x is the extension of a property that x doesn't exemplify. But we now show that the relation term $[\lambda x\, \exists G(x = \epsilon G \,\&\, \neg Gx)]$ is provably empty, and that the individual term $\epsilon[\lambda x \exists G(x=\epsilon G \& \neg Gx)]$ is as well.

(**331**) **Theorems:** Some Interesting Empty Terms. In our system, the expression $[\lambda x\, \exists G(x = \epsilon G \,\&\, \neg Gx)]$ is not a core λ-expression, as defined in (9.2); by the Encoding Formula Convention (17.3) and the definition of identity (23.1), the variable x in $x=\epsilon G$ occurs in encoding position (9.1) and so the λ in $[\lambda x \exists G(x= \epsilon G \,\&\, \neg Gx)]$ binds a variable that occurs in encoding position in the matrix.[212] So (39.2) doesn't assert $[\lambda x\, \exists G(x=\epsilon G \,\&\, \neg Gx)]\downarrow$. Instead, one can prove:

[212]The Encoding Formula Convention (17.3) requires us to regard the variable x in $x = \epsilon G$ as occurring free in encoding position since the *definiens* of $x=\epsilon G$ (23.1) is:

$$(O!x \,\&\, O!\epsilon G \,\&\, \Box\forall F(Fx \equiv F\epsilon G)) \vee (A!x \,\&\, A!\epsilon G \,\&\, \Box\forall F(xF \equiv \epsilon GF))$$

Here, the variable x occurs free in encoding position in the formula xF.

(.1) $\neg[\lambda x\, \exists G(x = \epsilon G \,\&\, \neg Gx)]\downarrow$

Thus, the λ-expression with the Fregean formula in (330) as matrix provably fails to be significant. Consequently, given how ϵF has been defined, it is provable that the extension of *being an extension x of a concept under which x doesn't fall* doesn't exist:

(.2) $\neg\epsilon[\lambda x\, \exists G(x = \epsilon G \,\&\, \neg Gx)]\downarrow$

Thus, the problematic extension provably doesn't exist.

(**332**) **Remark:** Digression on Frege's Definition of Set Membership. We mentioned in (316) that the definition of $y \in x$ is similar to Frege's definition, but differs from it in a number of ways. Those readers with an interest in a more exact comparison with Frege's work may therefore find the following to be of interest. In Frege 1893 (§34, [2013, 52]), Frege says:

> ...our concern is only to designate the value of the function $\Phi(\xi)$ for the argument Δ, that is, $\Phi(\Delta)$, using 'Δ' and '$\dot{\varepsilon}\Phi(\varepsilon)$'. I do so in this way:
>
> '$\Delta \frown \dot{\varepsilon}\Phi(\varepsilon)$'
>
> which is to be co-referential with '$\Phi(\Delta)$'.

So Frege doesn't really say that $\frown$ is the notion of membership in a set, or even membership in an extension. Immediately after the above passage, he offers the following, formal definition of the binary function $\xi \frown \zeta$, where we've replaced Frege's variable u with the variable x, and Frege's variable a with the variable y:

$$\Vdash \backslash\dot{\alpha}\left(\begin{array}{l} \top\!\overset{\mathfrak{g}}{\smile}\!\top\!\!\top\; \mathfrak{g}(y) = \alpha \\ \qquad\quad \llcorner\; x = \dot{\varepsilon}\mathfrak{g}(\varepsilon) \end{array}\right) = y \frown x \qquad (A$$

What's being defined here is the value of the function $y \frown x$. And Frege defines it as: *the* object α such that there exists a function $\mathfrak{g}$ for which x is the course-of-values for $\mathfrak{g}$ and α is the value of $\mathfrak{g}$ for the argument y. But let's limit the definition to the case where $\mathfrak{g}$ is a concept, so that we can replace it with our variable G. Then we might initially render Frege's definition in our notation as:

$$y \frown x \;=_{df}\; \imath\alpha\exists G(x = \dot{\varepsilon}G(\varepsilon) \,\&\, G(y) = \alpha)$$

However, in the case where G is a concept, the course-of-values x becomes the extension of G and the α in Frege's definiens ranges over truth values and, indeed, signifies the truth-value The True when y falls under the concept G. So the claim $G(y) = \alpha$ in Frege's definition can be represented by the exemplification predication Gy. That means the variable α is no longer needed and so we can transform Frege's definition-by-=, which assigns a value to the function term '$y \frown x$', into a definition-by-$\equiv$ that assigns truth conditions to the formula '$y \frown x$' – we eliminate the variable-binding description operator and the variable α from Frege's definiens, so that the definiens in Frege's definition now becomes:

$$y \frown x \ \equiv_{df} \ \exists G(x = \dot{\varepsilon}G(\varepsilon) \,\&\, Gy)$$

Finally, Frege's claim $x = \dot{\varepsilon}\mathfrak{g}(\varepsilon)$ can be represented in our notation as $x = \epsilon G$. However, by (327.3)★, $x = \epsilon G$ is materially equivalent to $\mathit{ExtensionOf}(x, G)$.[213] So Frege's definiens, in our notation, now becomes:

$$y \frown x \ \equiv_{df} \ \exists G(\mathit{ExtensionOf}(x, G) \,\&\, Gy)$$

This is the definition of $y \in x$ found in (316) above.

(**333**) ★**Corollary:** Fregean Principle of Extensions. It is an immediate, though not modally strict, corollary of (317.1) that an individual x is a member of the extension of F if and only if x exemplifies F:

$$x \in \epsilon F \equiv Fx$$

Cf. Frege 1893, §55, Theorem 1 (2013, 75), where Frege proves $f(a) = a \frown \dot{\varepsilon} f(\varepsilon)$.

(**334**) ★**Theorem:** The Extension of G is a Class. Finally, *the* extension of a property is a class:

$$\mathit{Class}(\epsilon G)$$

This captures a claim found in Carnap (1947, 19, 4-14), but as a logico-metaphysical claim about properties, not a semantic claim about predicates.

(**335**) **Remark:** Actual Extensions. As the reader works through the following definitions and theorems, it should be kept in mind that many of the theorems that fail to be modally strict could be turned into modally strict theorems by defining:

$$\mathit{ActualExtensionOf}(x, G) \ \equiv_{df} \ A!x \,\&\, \forall F(xF \equiv \mathcal{A}\forall z(Fz \equiv Gz))$$

[213] Since Frege didn't work within a modal context, it is not clear how $x = \dot{\varepsilon}\mathfrak{g}(\varepsilon)$ would or should behave when we evaluate it with respect to such a context; it depends on whether the term $\dot{\varepsilon}\mathfrak{g}(\varepsilon)$ is interpreted rigidly and whether identity claims are necessary if true. In our system, the counterpart of $x = \dot{\varepsilon}\mathfrak{g}(\varepsilon)$, namely, $x = \epsilon G$, is a rigid condition on properties: the claim $\Box\forall x(x = \epsilon G \rightarrow \Box x = \epsilon G)$ follows by GEN and RN from an instance of the necessity of identity. But (327.4)★, i.e., $\mathit{ExtensionOf}(x, G) \equiv x = \epsilon G$, is not a modally strict theorem, and so it is to be expected that $\mathit{ExtensionOf}(x, G)$ is not a rigid condition on properties. This should be relatively easy to see from the discussion in Remark (326).

It should be clear, then why we have substituted the condition $\mathit{ExtensionOf}(x, G)$ for the equivalent condition $x = \epsilon G$ in the Fregean definition of membership. This allows certain general theorems about extensions to be proved in a modally strict manner. By contrast, if we had preserved $x = \epsilon G$ as part of the definition, then since ϵG is defined in terms of a rigid definite description, certain general theorems about membership would fail to be modally strict.

This definition also yields as theorems:

$$\exists! x ActualExtensionOf(x, G)$$

$$\imath x ActualExtensionOf(x, G)\downarrow$$

Thus, we could introduce the notation:

$$\hat{\epsilon} G =_{df} \imath x ActualExtensionOf(x, G)$$

Clearly, $\hat{\epsilon} G$ is a *strictly* canonical object, since the application of GEN to an appropriate instance of axiom (46.1) yields a modally strict proof of:

$$\forall F(\mathcal{A}\forall z(Fz \equiv Gz) \rightarrow \Box\mathcal{A}\forall z(Fz \equiv Gz))$$

The reader should therefore keep in mind that non-modally strict theorems about extensions and ϵGcan often be converted to modally strict theorems about actual extensions and $\hat{\epsilon} G$ by strategic placement of the actuality operator $\mathcal{A}$.

Exercises: (a) Show that:

$$\forall x\forall G(ActualExtensionOf(x, G) \rightarrow \Box ActualExtensionOf(x, G))$$

(b) Use $ActualExtensionOf(x, G)$ to formulate modally-strict counterparts of the theorems proved above. (c) Show that if we define:

$$Class^*(x) \equiv_{df} \exists G(ActualExtensionOf(x, G))$$

then it follows that $\forall x(Class^*(x) \rightarrow \Box Class^*(x))$.

10.5 Interlude: Restricted Variables

Our next goal in the application of object theory is to derive the basic principles of natural classes (logical sets). But to simplify the expression of the theorems, we shall adopt the expedient of restricted variables. While the use of bound restricted variables to express theorems and free restricted variables to state definitions is straightforward enough, the presence of free restricted variables can easily lead one astray when reasoning. In this section, then, we'll discuss the principles and conventions that govern the use of restricted variables, and explain the problems of reasoning with free restricted variables in modal contexts.

(**336**) **Metadefinition:** Restricted Variables. A *restricted variable* is, intuitively, a variable introduced to range over just those individuals or relations satisfying a *restriction condition* ψ that meets three requirements:

(.1) ψ contains a single free, unrestricted variable, say α,

(.2) ψ is *strictly non-empty* in the sense that $\vdash_{\Box} \exists\alpha\psi$, and

(.3) ψ has *strict existential import* in the sense that $\vdash_{\Box} \psi^{\tau}_{\alpha} \to \tau\downarrow$, for any term τ substitutable for α.

For example, the conditions $O!x$, $D!x$, $\mathit{Propositional}(F)$, and $\mathit{Class}(x)$ are restriction conditions. Let's confirm this before we indicate how one might introduce corresponding restricted variables. By inspection, each condition contains a single free variable. Moreover, each is *strictly non-empty* since all of the following are provable by modally strict means: $\exists xO!x$, $\exists F(\mathit{Propositional}(F))$, and $\exists x\mathit{Class}(x)$ (exercises). And each has strict existential import, for the following are all modally strict theorems:

- $O!\kappa \to \kappa\downarrow$, for any individual term κ
  $D!\kappa \to \kappa\downarrow$, for any individual term κ
- $\mathit{Propositional}(\Pi) \to \Pi\downarrow$, for any property term Π
- $\mathit{Class}(\kappa) \to \kappa\downarrow$, for any individual term κ

We leave the first two as exercises, but to see that the third holds, assume $\mathit{Class}(\kappa)$. Then by definition (312.2), it follows that $\exists G(\mathit{ExtensionOf}(\kappa, G))$, at which point definition (312.1) implies $A!\kappa$, and thus, by axiom (39.5.a), $\kappa\downarrow$.[214]

Later, we'll discuss what may be labeled *weak* or *empty* restriction condition. A *weak restriction condition* is a condition ψ such that:

- ψ has a single free variable α
- ψ is *non-empty* in the sense that $\vdash \exists\alpha\psi$, and
- ψ has *existential import* in the sense that $\vdash \psi^{\tau}_{\alpha} \to \tau\downarrow$, for any term τ substitutable for α.

Speaking loosely, we might say that a weak restriction condition differs from a restriction condition in that modally strict proofs aren't required to show that ψ is non-empty and has existential import. Weak restriction conditions are of interest when $\exists\alpha\psi$ is either a modally fragile axiom or a $\star$-theorem.

By contrast, an *empty restriction condition* is a condition ψ that has a single free variable α and has existential import, but which may be *empty*, i.e., it is not a theorem that $\exists\alpha\psi$. One might want to introduce restricted variables for empty restriction conditions in situations where the condition ψ is well-defined and it is assumed (but not axiomatic) that $\exists\alpha\psi$, without this claim being a known theorem.

[214] By contrast, $\neg\mathit{Class}(x)$ is not a restriction condition. Though it contains a single free variable and is strictly non-empty (i.e., $\vdash_{\Box} \exists x\neg\mathit{Class}(x)$), it doesn't have existential import. The formula $\neg\mathit{Class}(\kappa)$ doesn't strictly imply $\kappa\downarrow$, for when $\neg\kappa\downarrow$, both $\neg\mathit{Class}(\kappa)$ and $\neg\kappa\downarrow$ hold.

In what follows, we'll avoid introducing variables for empty restriction conditions though, on occasion, we introduce variables for weak restriction conditions. We'll see some examples of weak restriction conditions when we consider the ordinary language properties *being a line* and *being a figure* and use them to develop the theory of *directions* and *shapes* (in Section 10.8.1), and when we discuss *stories* (in Section 12.6). When we introduce variables for directions, shapes, stories, etc., they will be *weak*, rather than *rigid*, restricted variables. See (342) for further discussion of empty and weak restriction conditions.

But let's now focus on restriction conditions as defined in (.1) – (.3) above. Since $O!x$, $D!x$, $Propositional(F)$, and $Class(x)$ are restriction conditions, we may, in each case, introduce distinguished variables that range over the entities satisfying the condition. Later, we'll use $u, v, \ldots$ to range over ordinary objects in certain contexts and to range over distinguished objects in other contexts, and use the letters $c, c', c'', \ldots$ to range over classes. Though we could also distinguish certain upper case letters to range over propositional properties, we shall not have the need to do so. By requiring that restriction conditions be strictly non-empty, we ensure that quantifier principles such as $\forall u\varphi \to \exists u\varphi$ and $\forall c\varphi \to \exists c\varphi$ hold, and we ensure that $u{\downarrow}$, $v{\downarrow}$, $c{\downarrow}$, and $c'{\downarrow}$ are still axiomatic. By requiring that restriction conditions have existential import, we ensure that definitions-by-equivalence using free restricted variables always abbreviate definitions in which the definiens will be false for any terms that either fail to denote or fail to meet the restriction condition. This will become clear in (338) below.

Restricted variables allow us to write long formulas and terms in *abbreviated* form. Thus, any expression (formula or term) written using a restricted variable is simply shorthand for an expanded expression that uses unrestricted variables. The primary purpose of restricted variables, therefore, is to reduce cognitive load by simplifying the expression of complex claims. Bound occurrences of restricted variables are easily eliminable; there is a straightforward way to interpret formulas containing them. Similarly, it is straightforward to interpret free occurrences of restricted variables in definitions. But, when reasoning, free occurrences of restricted variables aren't as easily eliminable; their interpretation is highly dependent on the context. The use of free restricted variables to state axioms and theorems, and when reasoning in proofs and derivations, is *highly problematic* when modal operators, actuality operators, and rigid definite descriptions are present; in particular, reasoning with free restricted variables often *increases* cognitive load and it is important to see why this is so, if only to counterbalance the strong temptation to use them. It is also worthwhile to determine conditions under which one may *safely* use them. We'll discuss these issues at length in (340) and (341).

Accordingly, our discussion of restricted variables will be divided into the following remarks:

- conventions for bound restricted variables (337),
- conventions for free restricted variables in definitions-by-≡ (338),
- conventions for free restricted variables in definitions-by-= (339),
- reasoning with bound restricted variables (340), and
- the problem of, and conditions for, using free restricted variables to assert axioms/theorems and to reason within derivations/proofs (341).

(**337**) **Remark:** Bound Restricted Variables. In this Remark, we consider restricted variables that are bound by sentence- and term-forming operators. We begin with a general statement of our conventions followed by a discussion of each convention and some of its implications. We can summarize these conventions in two groups – one group for the quantifiers and one group for the term-forming operators.

To introduce our conventions for the quantifiers $\forall$ and $\exists$, let:

- ψ be any restriction condition, as defined in (336),
- α be the unrestricted variable that occurs free in ψ (where α is either an individual variable or an n-ary relation variable, for some n),
- γ be introduced a variable of the same type as α but restricted by the condition ψ, and
- φ be any formula in which γ is substitutable for α.

Then we may introduce, as abbreviations, formulas in which restricted variables are bound by quantifiers:

(.1) $\forall\gamma\varphi^{\gamma}_{\alpha} =_{abbr} \forall\alpha(\psi \to \varphi)$

(.2) $\exists\gamma\varphi^{\gamma}_{\alpha} =_{abbr} \exists\alpha(\psi \,\&\, \varphi)$

Strictly speaking, since we only have one primitive quantifier in the system, we don't need the second abbreviation; we can *derive* from (.1) that $\exists\gamma\varphi^{\gamma}_{\alpha}$ abbreviates $\exists\alpha(\psi \,\&\, \varphi)$, by the definition of $\exists$.[215]

[215]To avoid clash of metavariables, we derive, for an arbitrary formula χ, that $\exists\gamma\chi^{\gamma}_{\alpha}$ abbreviates $\exists\alpha(\psi \,\&\, \chi)$. By definition (18.4), $\exists\gamma\chi^{\gamma}_{\alpha}$ expands to $\neg\forall\gamma\neg\chi^{\gamma}_{\alpha}$. Now let φ in (.1) be $\neg\chi$. Then, (.1) tells us that the universal claim in the scope of the initial negation operator, i.e., $\forall\gamma\neg\chi^{\gamma}_{\alpha}$, is an abbreviation of $\forall\alpha(\psi \to \neg\chi)$. So $\neg\forall\gamma\neg\chi^{\gamma}_{\alpha}$ abbreviates $\neg\forall\alpha(\psi \to \neg\chi)$, i.e., $\exists\alpha\neg(\psi \to \neg\chi)$. But by definition (18.1), we know $\psi \,\&\, \chi \equiv_{df} \neg(\psi \to \neg\chi)$. So by a Rule of Substitution (160.3), it follows that $\exists\alpha(\psi \,\&\, \chi)$.

Before we introduce our conventions for the term-forming operators $\imath$ and λ, it would serve well to give some examples of (.1) and (.2). A simple example of (.1) can be obtained by considering a theorem about discernible ($D!$) objects, defined in (273.2). Apply GEN to theorem (273.30) to obtain $\forall x(D!x \rightarrow x =_D x)$). Using convention (.1) however, with u as a restricted variable ranging over discernible objects, we may abbreviate this claim to:

$$\forall u(u =_D u)$$

An example that shows how (.1) suffices to interpret a string of universal quantifiers binding restricted variables is left to a footnote.[216] For a rather different example of (.1), let c be a restricted individual variable ranging over classes. Then we may write the (true) claim $\neg\forall x(Class(x) \rightarrow \exists y(y \in x))$ (i.e., not every class has members), as $\neg\forall c\exists y(y \in c)$. Moreover, we may write the (false) claim $\forall x(Class(x) \rightarrow \neg\exists y(y \in x))$ (i.e., every class fails to have members) more simply as $\forall c\neg\exists y(y \in c)$.

For an example of (.2), note that theorem (210.3) implies $\exists x(O!x \,\&\, \neg\mathcal{A}E!x)$. If we now let u be a restricted variable ranging over ordinary objects, we may therefore write this last claim as $\exists u\neg\mathcal{A}E!u$. Similarly, the (true) claim that some class fails to have members, i.e., $\exists x(Class(x) \,\&\, \neg\exists y(y \in x))$ may be more simply written with restricted variables as $\exists c\neg\exists y(y \in c)$. And, using both u and v as restricted variables ranging over ordinary objects, we leave it to the reader to show why $\forall u\exists v(v =_E u)$, by (.1) and (.2), abbreviates $\forall x(O!x \rightarrow \exists y(O!y \,\&\, y =_E x))$.

We now introduce our convention for the term-forming operators $\imath$ and λ. Let ψ, α, γ, and φ be as above but where ψ is a restriction condition on individuals, so that α and γ are some *individual* variables ν and μ, respectively, where μ is substitutable for ν in φ. Then we may introduce, as abbreviations, definite descriptions in which restricted variables are bound by the $\imath$ symbol:

(.3) $\imath\mu\varphi^{\mu}_{\nu} \;=_{abbr}\; \imath\nu(\psi \,\&\, \varphi)$

[216] Consider the following claim, which is easily derivable from theorem (273.7): $(D!x \,\&\, D!z) \rightarrow (\forall F(Fx \equiv Fz) \rightarrow x = z)$. By (88.7.a) and (88.7.b), this is equivalent to $D!x \rightarrow (D!z \rightarrow (\forall F(Fx \equiv Fz) \rightarrow x = z))$. After applying GEN to universally generalize on z, moving the quantifier $\forall z$ across the antecedent by (95.2), and then applying GEN to universally generalize on x, it follows that $\forall x(D!x \rightarrow \forall z(D!z \rightarrow (\forall F(Fx \equiv Fz) \rightarrow x = z)))$. Using convention (.1) however, with u, v as restricted variables over discernible objects, we may abbreviate the embedded universal claim and so obtain $\forall x(D!x \rightarrow \forall v(\forall F(Fx \equiv Fv) \rightarrow x = v))$, and by a second application of (.1), shorten this to $\forall u\forall v(\forall F(Fu \equiv Fv) \rightarrow u = v)$. It should be evident from this example that since the following two claims, $(\psi_1 \,\&\, \ldots \,\&\, \psi_n) \rightarrow \varphi$ and $\psi_1 \rightarrow (\ldots \rightarrow (\psi_n \rightarrow \varphi)\ldots)$, are equivalent, our system is indifferent as to which of the following $\forall\gamma_1 \ldots \gamma_n\varphi^{\gamma_1,\ldots,\gamma_n}_{\alpha_n,\ldots,\alpha_n}$ abbreviates:

$$\forall\alpha_1 \ldots \forall\alpha_n((\psi_1 \,\&\, \ldots \,\&\, \psi_n) \rightarrow \varphi)$$

$$\forall\alpha_1 \ldots \forall\alpha_n(\psi_1 \rightarrow (\ldots \rightarrow (\psi_n \rightarrow \varphi)\ldots))$$

As an example of (.3), note that $\imath u(u \neq u)$ could be used as shorthand for either of the empty descriptions $\imath x(O!x \,\&\, x \neq x)$ or $\imath x(D!x \,\&\, x \neq x)$, depending on whether u is a restricted variable for ordinary objects or discernible objects. For a more interesting example, suppose we could prove $\exists!x(Class(x) \,\&\, \neg\exists y(y \in x))$. It would follow that $\imath x(Class(x) \,\&\, \neg\exists y(y \in x))\downarrow$. Then we could abbreviate this last claim as $\imath c \neg\exists y(y \in c)\downarrow$.

Finally, to introduce our convention for the term-forming operator λ, let:

- $\psi_1, \ldots, \psi_n$ be any restriction conditions on individuals,
- $\nu_1, \ldots, \nu_n$ be, respectively, the unrestricted and distinct individual variables that occur free in $\psi_1, \ldots, \psi_n$,
- $\mu_1, \ldots, \mu_n$ be respectively introduced as distinct individual variables restricted by the conditions $\psi_1, \ldots, \psi_n$, and
- φ be any formula in which $\mu_1, \ldots, \mu_n$ are substitutable, respectively, for $\nu_1, \ldots, \nu_n$.

Then we may introduce, as abbreviations, λ-expressions in which restricted variables are bound by the λ symbol:

(.4) $[\lambda \mu_1 \ldots \mu_n \, \varphi^{\mu_1,\ldots,\mu_n}_{\nu_1,\ldots,\nu_n}] =_{abbr} [\lambda \nu_1 \ldots \nu_n \, \psi_1 \,\&\, \ldots \,\&\, \psi_n \,\&\, \varphi]$

For an example of (.4), consider that $[\lambda u \neg E!u]$ might abbreviate $[\lambda x \, O!x \,\&\, \neg E!x]$ or $[\lambda x \, D!x \,\&\, \neg E!x]$, depending on the context. And $[\lambda c \, \neg\exists y(y \in c)]$ abbreviates $[\lambda x \, Class(x) \,\&\, \neg\exists y(y \in x)]$. And if we let $\psi_1, \ldots, \psi_n$ be $D!x_1, \ldots, D!x_n$, respectively, let $u_1, \ldots, u_n$ be restricted variables ranging over discernible individuals, and let φ be $\neg R^n x_1 \ldots x_n$, then $[\lambda u_1 \ldots u_n \, \neg R^n u_1 \ldots u_n]$ abbreviates $[\lambda x_1 \ldots x_n \, D!x_1 \,\&\, \ldots \,\&\, D!x_n \,\&\, \neg Rx_1 \ldots x_n]$, the latter which has been established as significant in (273.15).

(**338**) **Remark:** Conventions for the Use of Free Restricted Variables in Definitions-by-Equivalence. One often finds free restricted variables in definitions-by-equivalence and definitions-by-identity. For example, set theorists introduce $\alpha, \beta, \gamma, \delta$ as restricted variables ranging over *ordinals*. They then define functions on, and properties of, ordinals by using these restricted variables.[217]

[217] In standard texts on set theory, one typically finds an *ordinal* to be defined as any set strictly well-ordered with respect to $\in$ and such that every element is also a subset. Then, using the restricted variables α and β to range over ordinals, one might find the following definition of the function term *the successor of* α and the notion α *is a limit ordinal*:

$$Suc(\alpha) =_{df} \alpha \cup \{\alpha\}$$

$$LimitOrdinal(\alpha) \equiv_{df} \alpha \neq \varnothing \,\&\, \neg\exists\beta(\alpha = Suc(\beta))$$

In Drake 1974 (41), for example, we find ordinal addition defined basically as follows, where λ ranges over limit ordinals and 0 is defined earlier (25) as $\varnothing$:

In this Remark, we confine ourselves to definitions-by-equivalence and begin with an example concerning classes.

Suppose we stipulate that a class c is *empty* just in case c has no members:

(A) $Empty(c) \equiv_{df} \neg\exists y(y \in c)$

How are we to understand this definition, i.e., what definition does (A) abbreviate? In the present text, the answer is that (A) abbreviates (B):

(B) $Empty(x) \equiv_{df} Class(x) \,\&\, \neg\exists y(y \in x)$

where (B) is governed by a version of Convention (17.2), i.e., one where the free variable x functions as a metavariable but the bound variable y needs no special interpretation (since we know that alphabetically-variant formulas are interderivable).[218] So (B) is, in turn, shorthand for the definition:

(B′) $Empty(\kappa) \equiv_{df} Class(\kappa) \,\&\, \neg\exists y(y \in \kappa)$, provided y doesn't occur free in κ

As we know from various remarks and theorems in the foregoing, (B′) extends our language to include formulas of the form $Empty(\kappa)$ and implies the closures of the following theorem schema:

(C) $Empty(\kappa) \equiv Class(\kappa) \,\&\, \neg\exists y(y \in \kappa)$, provided y doesn't occur free in κ.

But it is important to see why (A) should *not* be interpreted as a kind of conditional or contextual definition, i.e., as shorthand for adding $Empty(\kappa)$ to the language and stipulating the closures of the following as new axioms:

(D) $Class(\kappa) \to (Empty(\kappa) \equiv \neg\exists y(y \in \kappa))$, provided y doesn't occur free in κ

Note that (C) and (D) aren't equivalent; (C) implies (D) but not vice versa, since $\varphi \equiv (\psi \,\&\, \chi)$ implies $\psi \to (\varphi \equiv \chi)$, but the converse doesn't hold.[219]

There are two (related) reasons why (D) is a problematic way to eliminate the restricted variables in (A). The first is that (D) fails the *eliminability* criterion for definitions-by-equivalence: (D) doesn't generally define $Empty(\kappa)$ but

$\alpha + 0 = \alpha$

$\alpha + Suc(\beta) = Suc(\alpha + \beta)$

$\alpha + \lambda = \bigcup_{\delta<\lambda}(\alpha + \delta)$

Though Drake says how we should interpret restricted variables when they appear as bound variables (1974, 22, 26), he never indicates what is meant when they occur free in definitions.

[218] It is worth mentioning that in a properly formulated definition with free restricted variables, we won't need existence clauses in the definiens, since by (336.3), any condition $\psi(\alpha)$ that is used to introduce a restricted variable γ has to be such that, for any term τ substitutable for α, it is provable that $\psi^{\tau}_{\alpha} \to \tau\!\downarrow$. So we don't need existence clauses in definientia using free restricted variables, since a definiens $\varphi(\gamma)$ will imply $\tau\!\downarrow$ if φ^{τ}_{γ} holds.

[219] Theorem (88.8.i) establishes that $\varphi \equiv (\psi \,\&\, \chi)$ implies $\psi \to (\varphi \equiv \chi)$. To see that the converse doesn't hold, consider the scenario in which ψ is false but φ and χ are both true. Then $\psi \to (\varphi \equiv \chi)$ is true, by failure of the antecedent, but $\varphi \equiv (\psi \,\&\, \chi)$ is false.

rather defines it only when κ is a class.[220] The second, related problem is that if κ is an empty singular term, then (D) doesn't give one the means to prove $\neg Empty(\kappa)$, whereas (B) does.

Of course, some texts allow conditional/contextual definitions-by-$\equiv$, but in the present work, we prefer to avoid them.[221] Any condition needed for a definition-by-$\equiv$ can be built into the definiens. Thus, we consider it preferable to regard (A) as an abbreviation of (B). (B) defines $Empty(x)$ in all contexts, and has the virtue that if we know $Class(x)$, then by Rule $\equiv$S, the biconditionals that (B) gives rise to become equivalent to $Empty(x) \equiv \neg\exists y(y \in x)$. So in what follows, we'll use free restricted variables *in a definition-by-*$\equiv$ on the model of (A) and (B).

We may summarize this discussion of free restricted variables in definitions-by-$\equiv$ as follows. Let ψ be a restriction condition with α free and γ be the restricted variable, and let φ and χ be conditions on α. Furthermore, let us use the simple notation $\chi(\gamma)$ and $\varphi(\gamma)$, respectively, instead of χ^γ_α and φ^γ_α. Then:

(.1) A definition of the form $\chi(\gamma) \equiv_{df} \varphi(\gamma)$ abbreviates the definition:

$$\chi \equiv_{df} \psi \,\&\, \varphi,$$

where this latter definition is governed by Convention (17.2).

Thus, (A) and (B) are a simple example.[222] Of course, (.1) should be generalized to the case where the definiens and definiendum have multiple (free) restricted variables as arguments. Let $\psi_1,\ldots,\psi_n$ be conditions on the distinct variables $\alpha_1,\ldots,\alpha_n$, respectively; i.e., ψ_i is a condition on α_i, for $1 \leq i \leq n$, and α_i is the only variable that occurs free in ψ_i. Suppose further that $\gamma_1,\ldots,\gamma_n$ have been introduced as distinct restricted variables of the same type, respectively, as $\alpha_1,\ldots,\alpha_n$, so that γ_i $(1 \leq i \leq n)$ ranges over the entities such that ψ_i. Moreover, for simplicity, let $\chi(\gamma_1,\ldots,\gamma_n)$ and $\varphi(\gamma_1,\ldots,\gamma_n)$, respectively, abbreviate $\chi^{\gamma_1,\ldots,\gamma_n}_{\alpha_1,\ldots,\alpha_n}$ and $\varphi^{\gamma_1,\ldots,\gamma_n}_{\alpha_1,\ldots,\alpha_n}$. Then:

(.2) A definition of the form $\chi(\gamma_1,\ldots,\gamma_n) \equiv_{df} \varphi(\gamma_1,\ldots,\gamma_n)$ abbreviates the definition:

[220] This gives rise to a version of Frege's infamous 'Julius Caesar' problem. The problem is that (D) doesn't give us the means to prove whether or not Julius Caesar ('j') is *empty*, even given that $\neg Class(j)$.

[221] Note that §8.6 in Suppes 1957 is titled Conditional Definitions and the discussion (pp. 165–166) is primarily about the definition of the operation symbol of division (x/y). But, since there are no definite descriptions in the system Suppes is discussing, he can't define x/y outright using a definition-by-=. Instead he defines it using a conditional definition-by-$\equiv$ as follows: if $y \neq 0$, then $x/y = z \equiv x = y \cdot z$. It is thus clear from his discussion that Suppes allows conditional definitions-by-$\equiv$. But we shall not. However, we'll discuss a form of conditional definition-by-= in (366) and in the following Remark (367).

[222] In the case of (A) and (B), α is the variable x, ψ is the formula $Class(x)$, φ is the formula $\neg\exists y(y \in x)$, γ is c, and χ is the expression $Empty(x)$. Then the definiendum $\chi(\gamma)$ becomes $Empty(c)$ and the definiens $\varphi(\gamma)$ becomes $\neg\exists y(y \in c)$.

$$\chi \;\equiv_{df}\; \psi_1 \,\&\, \ldots \,\&\, \psi_n \,\&\, \varphi,$$

where this latter definition is governed by convention (17.2).

To see an example, consider the following, which is officially presented as item (350) in the next section. The definition for c is the *union of* c' and c'' is:

$$UnionOf(c,c',c'') \;\equiv_{df}\; \forall y(y \in c \equiv y \in c' \vee y \in c'')$$

This has the form:

$$\chi(\gamma_1,\gamma_2,\gamma_3) \;\equiv_{df}\; \varphi(\gamma_1,\gamma_2,\gamma_3)$$

and is to be regarded as shorthand for:

$$UnionOf(x,z,w) \;\equiv_{df}\; Class(x) \,\&\, Class(z) \,\&\, Class(w) \,\&\, \forall y(y \in x \equiv y \in z \vee y \in w)$$

which has the form:

$$\chi \;\equiv_{df}\; \psi_1 \,\&\, \psi_2 \,\&\, \psi_3 \,\&\, \varphi$$

where the substitutions for the metavariables are straightforward.[223] And by Convention (17.2), the strict form of the above definition is:

$$\begin{aligned}&UnionOf(\kappa_1,\kappa_2,\kappa_3) \;\equiv_{df}\\ &\quad Class(\kappa_1) \,\&\, Class(\kappa_2) \,\&\, Class(\kappa_3) \,\&\, \forall y(y \in \kappa_1 \equiv y \in \kappa_2 \vee y \in \kappa_3),\\ &\qquad \text{provided } y \text{ doesn't occur free in } \kappa_1,\ \kappa_2, \text{ or } \kappa_3\end{aligned}$$

One final observation is that our conventions, thus far, for bound and free restricted variables lead to some redundancies. For an example of such a redundancy, consider the following. In (344), we define a class c to be *empty* just in case it has no members:

$$Empty(c) \;\equiv_{df}\; \neg\exists y(y \in c) \tag{344}$$

Now suppose we want to prove $\exists c Empty(c)$. Then by our convention (337.2) for bound restricted variables, we have to show:

(ϑ) $\exists x(Class(x) \,\&\, Empty(x))$

But by our convention (.1), the definition of $Empty(c)$ (344), which we repeated above, abbreviates:

$$Empty(x) \;\equiv_{df}\; Class(x) \,\&\, \neg\exists y(y \in x)$$

By a Rule of Substitution (160.3), we can exchange the definiens and definiendum when they occur as subformulas. So to show (ϑ), it would suffice to show:

[223] Specifically, α_1 is x, α_2 is z and α_3 is w, ψ_1 is $Class(x)$, ψ_2 is $Class(z)$, ψ_3 is $Class(w)$, χ is $UnionOf(x,z,w)$, φ is $\forall y(y \in x \equiv y \in z \vee y \in w)$, and $\gamma_1,\gamma_2,\gamma_3$ are c,c',c'', respectively.

$\exists x(Class(x) \,\&\, Class(x) \,\&\, \neg\exists y(y \in x))$

But here, our conventions have lead us to prove something with an otiose conjunct. We can simply ignore this redundancy, for the idempotence of &, i.e., $(\varphi \,\&\, \varphi) \equiv \varphi$, is a modally strict theorem (85.6). So we may simplify and prove only:

(ξ) $\exists x(Class(x) \,\&\, \neg\exists y(y \in x))$

We'll henceforth disregard other redundancies of this kind that might arise from combining our conventions for restricted variables.

(**339**) **Remark:** On the Use of Free Restricted Variables in Definitions-by-Identity. The conventions for interpreting free restricted variables in definitions-by-identity are straightforward. Let $\psi_1, \ldots, \psi_n$, $\alpha_1, \ldots, \alpha_n$, and $\gamma_1, \ldots, \gamma_n$ be as stipulated in the foregoing remarks, but where $\alpha_1, \ldots, \alpha_n$ occur free in both $\iota\nu\varphi$ and $[\lambda\nu_1 \ldots \nu_n\, \varphi]$. Finally, let κ and Π be definienda in which $\alpha_1, \ldots, \alpha_n$ occur free, and let:

- $\kappa(\gamma_1, \ldots, \gamma_n)$ represent $\kappa^{\gamma_1,\ldots,\gamma_n}_{\alpha_1,\ldots,\alpha_n}$
- $\Pi(\gamma_1, \ldots, \gamma_n)$ represent $\Pi^{\gamma_1,\ldots,\gamma_n}_{\alpha_1,\ldots,\alpha_n}$
- $\varphi(\gamma_1, \ldots, \gamma_n)$ represent $\varphi^{\gamma_1,\ldots,\gamma_n}_{\alpha_1,\ldots,\alpha_n}$

Then we observe the following conventions:

(.1) $\kappa(\gamma_1, \ldots, \gamma_n) =_{df} \iota\nu\varphi(\gamma_1, \ldots, \gamma_n)$ abbreviates the definition:

$$\kappa =_{df} \iota\nu(\psi_1 \,\&\, \ldots \,\&\, \psi_n \,\&\, \varphi)$$

(.2) $\Pi(\gamma_1, \ldots, \gamma_n) =_{df} [\lambda\nu_1 \ldots \nu_n\, \varphi(\gamma_1, \ldots, \gamma_n)]$ abbreviates the definition:

$$\Pi =_{df} [\lambda\nu_1 \ldots \nu_n\, \psi_1 \,\&\, \ldots \,\&\, \psi_n \,\&\, \varphi]$$

Let's start with an example of (.2). Let u be a restricted variable ranging over discernible objects. Then, if given a property F, we might define the property F^{+u} as the property *being F or identical*$_D$ *to u*, as follows:

$$F^{+u} =_{df} [\lambda x\, Fx \,\&\, x =_D u]$$

Then (.2) tells us that this abbreviates the definition:

$$F^{+y} =_{df} [\lambda x\, D!y \,\&\, Fx \,\&\, x =_D y]$$

In this case, α is y, γ is u, and ν is x, so that:

- Π is F^{+y},
- $[\lambda\nu\, \psi \,\&\, \varphi]$ is $[\lambda x\, D!y \,\&\, Fx \,\&\, x =_D y]$,

- $\Pi(\gamma)$ is F^{+u}, and
- $[\lambda\nu\, \varphi(\gamma)]$ is $[\lambda x\, Fx \,\&\, x =_D u]$

As an example of (.1), let γ_1 be u and γ_2 be v, where u and v are restricted variables ranging over discernible ($D!$) objects, as defined in (273.2). We might then define *the natural pair class* of discernible objects u and v, written $\{u,v\}_D$ as follows:

(A) $\{u,v\}_D =_{df} \imath c\forall y(y \in c \equiv y =_D u \vee y =_D v))$

In fact, we'll define natural pair classes somewhat differently in (381), but for now, (A) offers a good illustration. Where ψ_1 is $D!z$ and ψ_2 is $D!w$, (.1) tells us that (A) abbreviates:

(B) $\{z,w\}_D =_{df} \imath c(D!z \,\&\, D!w \,\&\, \forall y(y \in c \equiv (y =_D z \vee y =_D w)))$

And if we then eliminate the bound restricted variable in the definiens, (B) abbreviates:

(C) $\{z,w\}_D =_{df} \imath x(Class(x) \,\&\, D!z \,\&\, D!w \,\&\, \forall y(y \in x \equiv (y =_D z \vee y =_D w)))$

Given that (A) and (B) abbreviate (C), the expression $\{\kappa,\kappa'\}_D$ is well-formed for any individual terms κ and κ', by our convention that variables in definitions function as metavariables (17.2.a). But we can't assume that $\{\kappa,\kappa'\}_D$ is significant for every pair of individual terms κ and κ'. For example, if it is known, by proof or by hypothesis, that one or both of κ and κ' is an empty term or that one or both of κ and κ' denotes an indiscernible ($\overline{D!}$) object, then the Rule of Definition by Identity (73) yields that $\{\kappa,\kappa'\}_D$ is an empty term, by the following extended argument.

Without loss of generality, consider just κ and assume $\neg\kappa\!\downarrow \vee \overline{D!}\kappa$. Then, in either case, $\neg D!\kappa$ (exercise). Moreover, in either case, $\Box\neg D!\kappa$.[224] Note that from $\neg D!\kappa$ it follows that no object satisfies the condition: x is a class such that both κ and κ' are discernible and such that the members of x are identical$_D$ to either κ or κ', i.e.,

(D) $\neg D!\kappa \vdash \neg\exists x(Class(x) \,\&\, D!\kappa \,\&\, D!\kappa' \,\&\, \forall y(y \in x \equiv (y =_D \kappa \vee y =_D \kappa')))$

A fortiori, $\neg D!\kappa$ implies that there is no *unique* such object, i.e.,

[224] We reason separately from the disjuncts $\neg\kappa\!\downarrow$ and $\overline{D!}\kappa$. If $\neg\kappa\!\downarrow$, then by (169.3), $\Box\neg\kappa\!\downarrow$. But $\neg\kappa\!\downarrow \to \neg D!\kappa$ is the contrapositive of axiom (39.5.a) and so by Rule RN, $\Box(\neg\kappa\!\downarrow \to \neg D!\kappa)$. The K axiom then implies $\Box\neg\kappa\!\downarrow \to \Box\neg D!\kappa$. Since we've established $\Box\neg\kappa\!\downarrow$, it follows that $\Box\neg D!\kappa$.

Alternatively, suppose $\overline{D!}\kappa$. Then $\kappa\!\downarrow$. We may use this fact and the fact that $[\lambda x\, \neg D!x]\!\downarrow$ (exercise) to derive, by β-Conversion, $[\lambda x\, \neg D!x]\kappa \equiv \neg D!\kappa$. So by definition of $\overline{D!}$ (273.2), it is a modally strict theorem that $\overline{D!}\kappa \equiv \neg D!\kappa$. Hence $\neg D!\kappa$, and so it follows by modus tollens from an appropriate instance of the T schema that $\neg\Box D!\kappa$. Now if we apply Rule RN and GEN to (273.8) and instantiate to κ, we obtain $\Box(D!\kappa \to \Box D!\kappa)$. Hence, by (172.3) $\neg\Box D!\kappa \equiv \Box\neg D!\kappa$. So $\Box\neg D!\kappa$.

(E) $\neg D!\kappa \vdash \neg\exists!x(Class(x) \,\&\, D!\kappa \,\&\, D!\kappa' \,\&\, \forall y(y \in x \equiv (y =_D \kappa \vee y =_D \kappa')))$

So by the Rule of Actualization:

(F) $\mathcal{A}\neg D!\kappa \vdash \mathcal{A}\neg\exists x(Class(x) \,\&\, D!\kappa \,\&\, D!\kappa' \,\&\, \forall y(y \in x \equiv (y =_D \kappa \vee y =_D \kappa')))$

But, from the previously noted fact that $\Box\neg D!\kappa$, it follows that $\mathcal{A}\neg D!\kappa$, by (132). It follows from this and (F) that:

$$\mathcal{A}\neg\exists x(Class(x) \,\&\, D!\kappa \,\&\, D!\kappa' \,\&\, \forall y(y \in x \equiv (y =_D \kappa \vee y =_D \kappa')))$$

Then by axiom (44.1), we have:

$$\neg\mathcal{A}\exists!x(Class(x) \,\&\, D!\kappa \,\&\, D!\kappa' \,\&\, \forall y(y \in x \equiv (y =_D \kappa \vee y =_D \kappa')))$$

So by theorem (176.2):

(G) $\neg\imath x(Class(x) \,\&\, D!\kappa \,\&\, D!\kappa' \,\&\, \forall y(y \in x \equiv (y =_D \kappa \vee y =_D \kappa')))\!\downarrow$

But (C) tells us, by the Rule of Definition by Identity (73), that:

(H) $(\imath x(Class(x) \,\&\, D!\kappa \,\&\, D!\kappa' \,\&\, \forall y(y \in x \equiv (y =_D \kappa \vee y =_D \kappa')))\!\downarrow \to$
$\{\kappa, \kappa'\}_D = \imath x(Class(x) \,\&\, D!\kappa \,\&\, D!\kappa' \,\&\, \forall y(y \in x \equiv (y =_D \kappa \vee y =_D \kappa')))) \,\&$
$(\neg\imath x(Class(x) \,\&\, D!\kappa \,\&\, D!\kappa' \,\&\, \forall y(y \in x \equiv (y =_D \kappa \vee y =_D \kappa')))\!\downarrow \to$
$\neg\{\kappa, \kappa'\}_D\!\downarrow)$

So (G) and the second conjunct of (H) imply $\neg\{\kappa, \kappa'\}_D\!\downarrow$.

Thus the inferential role and conventions for using restricted free variables in definitions-by-identity may combine to preserve the *garbage in*, *garbage out* principle. When we substitute an argument term (of the appropriate type) for a free *restricted* variable in a term-forming operator with such a variable, the resulting complex term fails to be significant whenever the argument term is empty or has a denotation not in the domain over which the restricted variable ranges. Our definitions-by-identity that have free restricted variables follow this pattern.

(**340**) **Remark:** (The Pitfalls of) Reasoning with Bound Restricted Variables. In this Remark, we examine the use of bound restricted variables when reasoning and, after examining some expected results, outline both the issues that arise and the precautions that have to be taken when modal, actuality, and description operators are involved. In what follows, assume that the formulas φ being discussed have only the bound restricted variables indicated in the example, and no free restricted variables.

It is well-known that if ψ is a restriction condition with α free and γ is a restricted variable introduced via this condition, then $\forall\alpha\varphi$ implies $\forall\gamma\varphi^{\gamma}_{\alpha}$, i.e., if a claim is true of everything in the domain, it is true for the restricted class. This follows from the principle $\varphi \to (\psi \to \varphi)$ (38.1) and we leave the proof to a footnote.[225] For example, if every individual is ordinary or abstract, i.e., $\forall x(O!x \vee A!x)$, then every class is ordinary or abstract, i.e., $\forall c(O!c \vee A!c)$. But, clearly, $\forall\gamma\varphi^{\gamma}_{\alpha}$ doesn't imply $\forall\alpha\varphi$; from the fact that every class is abstract ($\forall cA!c$), it doesn't follow that every object is abstract ($\forall xA!x$).

Dual principles hold for the restricted variables bound by existential quantifiers. A theorem of the form $\exists\gamma\varphi^{\gamma}_{\alpha}$ implies $\exists\alpha\varphi$,[226] but again the converse doesn't hold. We leave it to the reader to construct examples.

Another consequence of our definition of a *restriction* condition was previously mentioned, namely, the fact that $\forall\gamma\varphi^{\gamma}_{\alpha} \to \exists\gamma\varphi^{\gamma}_{\alpha}$ is a modally strict theorem.[227] Later, in Remark (342), we'll discuss (a) empty restriction conditions ψ (i.e., ones that nothing satisfies), for which $\forall\gamma\varphi^{\gamma}_{\alpha} \to \exists\gamma\varphi^{\gamma}_{\alpha}$ doesn't hold, and (b) weak restriction conditions ψ (i.e., ones that are satisfied but not by a modally strict proof), for which $\forall\gamma\varphi^{\gamma}_{\alpha} \to \exists\gamma\varphi^{\gamma}_{\alpha}$ holds but not as a modally strict theorem.

It is also important to observe that adjacent pairs of universal quantifiers commute even if one member of the pair binds a restricted variable and the other binds an unrestricted variable. To be specific, the following is a modally strict fact:

> $\forall\gamma\forall\beta\varphi^{\gamma}_{\alpha} \equiv \forall\beta\forall\gamma\varphi^{\gamma}_{\alpha}$, where γ is a restricted variable introduced by some restriction condition which has only α free

The proof is left to a footnote.[228]

Though the foregoing principles for bound restricted variables suggest that reasoning with them is unproblematic, it is important to note that *modal* principles involving such variables *fail*. For example, given our conventions in (337.1) and (337.2), the Barcan and Converse Barcan Formulas fail for quantified claims with restricted variables. The following are *not* theorems:

---

[225] Assume $\forall\alpha\varphi$. Then, φ, by $\forall$E. Hence, $\psi \to \varphi$, by axiom (38.1). So by GEN, $\forall\alpha(\psi \to \varphi)$. Hence, by our convention (337.1), it follows that $\forall\gamma\varphi^{\gamma}_{\alpha}$.

[226] Assume $\exists\gamma\varphi^{\gamma}_{\alpha}$. Then by (.2), $\exists\alpha(\psi \,\&\, \varphi)$. Hence, $\exists\alpha\varphi$.

[227] Assume $\forall\gamma\varphi^{\gamma}_{\alpha}$. Then by (.1) this abbreviates $\forall\alpha(\psi \to \varphi)$. But, by hypothesis, ψ is a restriction condition, and so $\exists\alpha\psi$ is a modally strict theorem, by (336). Let τ be any witness, so that we know ψ^{τ}_{α}. Then φ^{τ}_{α}, and by conjoining the last two claims and using our rules for the existential quantifier, it follows that $\exists\alpha(\psi \,\&\, \varphi)$.

[228] Suppose ψ is the restriction condition. Then if we eliminate the bound restricted variable from both sides of the biconditonal, we have to show:

$$\forall\alpha(\psi \to \forall\beta\varphi) \equiv \forall\beta\forall\alpha(\psi \to \varphi)$$

($\to$) Assume $\forall\alpha(\psi \to \forall\beta\varphi)$. Then $\psi \to \forall\beta\varphi$, by $\forall$E. We want to show $\forall\beta\forall\alpha(\psi \to \varphi)$. Since α and β don't occur free in our assumption, it suffices by GEN to show $\psi \to \varphi$. So assume ψ. Then $\forall\beta\varphi$ and hence φ. ($\leftarrow$) Assume $\forall\beta\forall\alpha(\psi \to \varphi)$. We want to show $\forall\alpha(\psi \to \forall\beta\varphi)$. Since α doesn't occur free in our assumption, it suffices to show $\psi \to \forall\beta\varphi$. So assume ψ. Since β doesn't occur free in our assumption, it suffices by GEN to show φ. But our initial assumption implies $\psi \to \varphi$, and ψ holds by assumption.

$$\forall\gamma\Box\varphi^{\gamma}_{\alpha} \rightarrow \Box\forall\gamma\varphi^{\gamma}_{\alpha}$$

$$\Box\forall\gamma\varphi^{\gamma}_{\alpha} \rightarrow \forall\gamma\Box\varphi^{\gamma}_{\alpha}$$

To see why, consider that the above abbreviate, respectively:

$$\forall\alpha(\psi \rightarrow \Box\varphi) \rightarrow \Box\forall\alpha(\psi \rightarrow \varphi)$$

$$\Box\forall\alpha(\psi \rightarrow \varphi) \rightarrow \forall\alpha(\psi \rightarrow \Box\varphi)$$

By appealing, for the sake of argument, to semantically-primitive possible worlds, one can see, intuitively, why these fail; we leave this to a footnote.[229] So when reasoning with a restricted variable γ, one may not commute $\forall\gamma$ and $\Box$, and the same applies to $\Diamond$ and $\exists\gamma$. This suggests that any quantified modal theorems derived from the Barcan and Converse Barcan formulas won't have instances in which the quantifier binds restricted variables.

There is a special case, however, where $\forall\gamma$ and $\Box$ do validly commute. If γ ranges over the entities that satisfy a *rigid* restriction condition, i.e., a restriction condition that, by a modally strict proof, is necessarily satisfied by anything that satisfies it, then we may safely apply RN or RA. In such a situation, the Barcan and Converse Barcan formulas are valid.

To see this, first recall that in (260.1), we stipulated that ψ is a *rigid* condition on α just in case $\vdash_{\Box} \forall\alpha(\psi \rightarrow \Box\psi)$. Then let us first define:

> ψ is a *rigid restriction condition on α* just in case ψ is a restriction condition on α that is also a rigid condition on α.

For example, $O!x$ is a rigid restriction condition on individuals, given that by GEN, it follows from (180.1) that $\vdash_{\Box} \forall x(O!x \rightarrow \Box O!x)$. And similarly for $D!x$, by

[229] In what follows, we simplify notation by writing $\psi^{\boldsymbol{e}}_{\alpha}$ as $\psi(\boldsymbol{e})$. To see that the first is false, consider any 2-world model in which $\boldsymbol{w}_0$ is the actual world and where $\boldsymbol{e}$ ranges over the entities in the range of the variable α:

at $\boldsymbol{w}_0$: $\forall\boldsymbol{e}\neg\psi(\boldsymbol{e})$

at $\boldsymbol{w}_1$: $\exists\boldsymbol{e}(\psi(\boldsymbol{e}) \,\&\, \neg\varphi(\boldsymbol{e}))$

Then the antecedent $\forall\alpha(\psi \rightarrow \Box\varphi)$ is true (i.e., true at $\boldsymbol{w}_0$) by failure of the antecedent, but the consequent $\Box\forall\alpha(\psi \rightarrow \varphi)$ is false (i.e., false at $\boldsymbol{w}_0$) because $\boldsymbol{w}_1$ is a world where $\forall\alpha(\psi \rightarrow \varphi)$ is false.

To see that the second is false, consider the following 2-world model in which entity $\boldsymbol{e}_1$ is in the range of the variable α:

at $\boldsymbol{w}_0$: $\psi(\boldsymbol{e}_1)$, $\varphi(\boldsymbol{e}_1)$, $\forall\boldsymbol{e}(\boldsymbol{e} \neq \boldsymbol{e}_1 \rightarrow \neg\psi(\boldsymbol{e}))$

at $\boldsymbol{w}_1$: $\forall\boldsymbol{e}\neg\psi(\boldsymbol{e})$, $\neg\varphi(\boldsymbol{e}_1)$

Then the antecedent $\Box\forall\alpha(\psi \rightarrow \varphi)$ is true (i.e., true at $\boldsymbol{w}_0$) because at both worlds, the claim $\forall\alpha(\psi \rightarrow \varphi)$ is true: at $\boldsymbol{w}_0$, if the value of α is $\boldsymbol{e}_1$, then both ψ and φ are true, so that $\psi \rightarrow \varphi$ is true, and if the value of α is something other than $\boldsymbol{e}_1$, then the antecedent is false, making $\psi \rightarrow \varphi$ true; and at $\boldsymbol{w}_1$, ψ is false for all values of α and so $\forall\alpha(\psi \rightarrow \varphi)$ is true. But the consequent $\forall\alpha(\psi \rightarrow \Box\varphi)$ is false (i.e., false at $\boldsymbol{w}_0$) because when α has the value $\boldsymbol{e}_1$, ψ is true at $\boldsymbol{w}_0$ while $\Box\varphi$ is false at $\boldsymbol{w}_0$ (since at $\boldsymbol{w}_1$, $\neg\varphi(\boldsymbol{e}_1)$ is false).

(273.8). By contrast, *Class*(x) is not a rigid restriction condition on individuals, given theorem (320.2); indeed, in (320.4) we saw that classes fail to be classes necessarily.[230]

Now we may say:

> γ is a *rigid restricted variable* just in case, for some rigid restriction condition ψ on α, γ is (introduced as) a variable whose values may be any α such that ψ.

So, we may now establish that if γ is a rigid restricted variable, then the Barcan and Converse Formulas hold with respect to γ, i.e., that $\forall\gamma\Box\varphi^{\gamma}_{\alpha} \to \Box\forall\gamma\varphi^{\gamma}_{\alpha}$ and that $\Box\forall\gamma\varphi^{\gamma}_{\alpha} \to \forall\gamma\Box\varphi^{\gamma}_{\alpha}$. It suffices to show that the formulas that they abbreviate are modally strict theorems:

- $\forall\alpha(\psi \to \Box\varphi) \to \Box\forall\alpha(\psi \to \varphi)$, when ψ is a rigid restriction condition.

  Proof. Assume:

  (ϑ) $\forall\alpha(\psi \to \Box\varphi)$

  Since we want to show $\Box\forall\alpha(\psi \to \varphi)$, it suffices, by BF, to show $\forall\alpha\Box(\psi \to \varphi)$, and since α isn't free in our assumption, it suffices to show $\Box(\psi \to \varphi)$. Note first that since ψ is, by hypothesis, a rigid restriction condition, we know $\vdash_{\Box} \forall\alpha(\psi \to \Box\psi)$. So by RN, $\Box\forall\alpha(\psi \to \Box\psi)$. By CBF, $\forall\alpha\Box(\psi \to \Box\psi)$, and so by $\forall$E, $\Box(\psi \to \Box\psi)$. But (ϑ) similarly implies $\psi \to \Box\varphi$. So by (172.6), $\Box(\psi \to \varphi)$, which is what we had to show.

- $\Box\forall\alpha(\psi \to \varphi) \to \forall\alpha(\psi \to \Box\varphi)$, when ψ is a rigid restriction condition.

  Proof. Assume $\Box\forall\alpha(\psi \to \varphi)$. To show $\forall\alpha(\psi \to \Box\varphi)$, it suffices, by GEN, to show $\psi \to \Box\varphi$, since α isn't free in our assumption. So assume ψ. Since ψ is a rigid restriction condition, we know that it is a modally strict theorem that $\forall\alpha(\psi \to \Box\psi)$. By $\forall$E, $\psi \to \Box\psi$. Hence $\Box\psi$. But by the Converse Barcan Formula (167.2), our assumption implies $\forall\alpha\Box(\psi \to \varphi)$. So, $\Box(\psi \to \varphi)$. It follows by the K axiom that $\Box\psi \to \Box\varphi$. Hence, $\Box\varphi$.

[230]Note that in (262), we saw three examples of rigid conditions on properties, namely, $\Box Fy$, $F = G$, and $\Box\forall y(Gy \to Fy)$. But these aren't (rigid) *restriction* conditions on properties, since they contain two free variables. If we assign y and G a value, say a and P (i.e., a and P are primitive constants, so that we know $a{\downarrow}$ and $P{\downarrow}$), respectively, then $\Box Fa$ and $F = P$ are rigid restriction conditions on properties, but $\Box\forall y(Py \to Fy)$ is not. (**Exercise:** Identify the clause in the definition of restriction condition that $\Box\forall y(Py \to Fy)$ fails and say why.) We've also seen two examples of *defined* restriction conditions: *Null*(x) is a restriction condition on objects and *Propositional*(F) is a restriction condition on properties. Theorems (266.1) and (276.4), by GEN, imply, respectively, that $\forall x(Null(x) \to \Box Null(x))$ and $\forall F(Propositional(F) \to \Box Propositional(F))$ are modally strict theorems.

Thus, the Barcan formulas apply to bound restricted variables when those variables are rigid.

A final observation about reasoning with conventions (337.1) and (337.2) is that there is an analogous phenomena affecting the actuality operator; both directions of axiom (44.3) fail to be necessary truths when the variables are restricted. The following are *not* modally strict theorems:

$$\forall\gamma\mathcal{A}\varphi^{\gamma}_{\alpha} \to \mathcal{A}\forall\gamma\varphi^{\gamma}_{\alpha}$$

$$\mathcal{A}\forall\gamma\varphi^{\gamma}_{\alpha} \to \forall\gamma\mathcal{A}\varphi^{\gamma}_{\alpha}$$

To see why, consider what they abbreviate:

$$\forall\alpha(\psi \to \mathcal{A}\varphi) \to \mathcal{A}\forall\alpha(\psi \to \varphi)$$

$$\mathcal{A}\forall\alpha(\psi \to \varphi) \to \forall\alpha(\psi \to \mathcal{A}\varphi)$$

Both of these latter *are* theorems, but they are not modally strict theorems. Again, we leave the argument to a footnote.[231] So it is important to recognize when reasoning with restricted variables that $\forall\gamma$ and $\mathcal{A}$ commute at the price of modal strictness.

Here again, though, the requirements that ψ be a rigid restriction condition and γ a rigid restricted variable restores the necessity of these commutation principles. For one can show that if ψ is a rigid restriction condition, then $\forall\gamma\mathcal{A}\varphi^{\gamma}_{\alpha} \to \mathcal{A}\forall\gamma\varphi^{\gamma}_{\alpha}$ and $\mathcal{A}\forall\gamma\varphi^{\gamma}_{\alpha} \to \forall\gamma\mathcal{A}\varphi^{\gamma}_{\alpha}$ abbreviate modally strict theorems,

[231] Here, we not only have to show that they are in fact $\star$-theorems, but also that their necessitations are invalid. To see that the first is a $\star$-theorem, assume $\forall\alpha(\psi \to \mathcal{A}\varphi)$. Then by theorem (130.1)$\star$, it suffices to show $\forall\alpha(\psi \to \varphi)$, and by GEN it suffices to show $\psi \to \varphi$. So assume ψ. But our initial assumption implies $\psi \to \mathcal{A}\varphi$, by $\forall$E. So $\mathcal{A}\varphi$. Hence, by (43)$\star$, φ. To see that the second is a theorem, assume $\mathcal{A}\forall\alpha(\psi \to \varphi)$. Then by GEN, it suffices to show $\psi \to \mathcal{A}\varphi$. So assume ψ. Our initial assumption implies, by (43)$\star$, that $\forall\alpha(\psi \to \varphi)$. So $\psi \to \varphi$ and, hence, φ. And so by (130.1)$\star$, $\mathcal{A}\varphi$.

Now to see that their necessitations are invalid, we again appeal to semantically primitive possible worlds. We just need to describe a model where these claims are true at $\boldsymbol{w}_0$ but fail at some other possible world $\boldsymbol{w}_1$. For the first, consider a model where:

at $\boldsymbol{w}_0$: $\exists\boldsymbol{e}(\psi(\boldsymbol{e}) \,\&\, \neg\varphi(\boldsymbol{e}))$

at $\boldsymbol{w}_1$: $\neg\exists\boldsymbol{e}\psi(\boldsymbol{e})$

Then the claim, $\forall\alpha(\psi \to \mathcal{A}\varphi) \to \mathcal{A}\forall\alpha(\psi \to \varphi)$, is true at $\boldsymbol{w}_0$ (because the antecedent $\forall\alpha(\psi \to \mathcal{A}\varphi)$ is false at $\boldsymbol{w}_0$), but false at $\boldsymbol{w}_1$ (because the antecedent $\forall\alpha(\psi \to \mathcal{A}\varphi)$ is true at $\boldsymbol{w}_1$ and the consequent $\mathcal{A}\forall\alpha(\psi \to \varphi)$ false at $\boldsymbol{w}_1$).

For the second, consider a model where:

at $\boldsymbol{w}_0$: $\neg\varphi(\boldsymbol{e}_1)$, $\neg\exists\boldsymbol{e}\psi(\boldsymbol{e})$

at $\boldsymbol{w}_1$: $\psi(\boldsymbol{e}_1)$

Then then the claim $\mathcal{A}\forall\alpha(\psi \to \varphi) \to \forall\alpha(\psi \to \mathcal{A}\varphi)$ is true at $\boldsymbol{w}_0$ (because the antecedent and consequent are both true at $\boldsymbol{w}_0$), but false at $\boldsymbol{w}_1$ (because the antecedent $\mathcal{A}\forall\alpha(\psi \to \varphi)$ is true at $\boldsymbol{w}_1$ and the consequent $\forall\alpha(\psi \to \mathcal{A}\varphi)$ is false at $\boldsymbol{w}_1$).

by arguments somewhat similar to the proofs displayed in the bulleted points above. We leave the proof of the first to a footnote, and leave the proof of the second to the reader.[232]

(**341**) **Remark and Metatheorem:** The Problem of Free Restricted Variables in Axioms/Theorems and Derivations/Proofs. Once bound restricted variables are used to assert axioms and theorems, and free restricted variables are used to formulate definitions, it won't be long before one attempts to *reason* with free restricted variables. The discussion in (340) already suggests why there may be concerns about doing so, but let's now explicitly identify some of the problems.

First, note that the use of free restricted variables immediately raises questions of interpretation. To see why, consider the formula $\neg\exists FuF$, in which u is a free restricted variable ranging over ordinary objects. It could be used to assert a theorem or a conclusion of a derivation, in which case it would be interpreted *conditionally*, as $O!x \to \neg\exists FxF$; indeed, on this understanding, $\neg\exists FuF$ would simplify the statement of the axiom for encoding (51). However, if one is *reasoning* in a derivation and one takes the formula $\neg\exists FuF$ as an assumption or premise, then it would be interpreted *conjunctively*, as $O!x \,\&\, \neg\exists FxF$, since $\neg\exists FuF$ carries the information that u is an ordinary object. So formulas with free restricted variables give rise to ambiguities, which can often be resolved by the context.

Despite the threat of ambiguity, the tendency to use free restricted variables is a strong one. Given that the conditional interpretation helps to simplify the statement of theorems and conclusions of reasoning, and the conjunctive interpretation helps to simplify assumptions and premises, one is tempted to let the context determine the interpretation of a formula with free restricted variables. But even if the context is always clear, the use of such variables can give rise to errors in reasoning, particularly in modal contexts.

We can approach the problem by first examining a simple case where we escape error only by a happy coincidence. Clearly, it is a modally strict theorem that $O!x \to O!x$. So, if we use free restricted variables, we should be able to assert this as the modally strict theorem $O!u$. Since the theorem is modally

[232]We want to show that what $\forall\gamma\mathcal{A}\varphi^{\gamma}_{\alpha} \to \mathcal{A}\forall\gamma\varphi^{\gamma}_{\alpha}$ abbreviates, namely:

$$\forall\alpha(\psi \to \mathcal{A}\varphi) \to \mathcal{A}\forall\alpha(\psi \to \varphi), \text{ where } \psi \text{ is a rigid restriction condition}$$

is a *modally strict* theorem. So assume:

$(\vartheta)\ \ \forall\alpha(\psi \to \mathcal{A}\varphi)$

Since we want to show $\mathcal{A}\forall\alpha(\psi \to \varphi)$, it suffices by axiom (44.3) to show $\forall\alpha\mathcal{A}(\psi \to \varphi)$. Since α isn't free in our assumption, it suffices by GEN to show $\mathcal{A}(\psi \to \varphi)$. Now since ψ is, by hypothesis, a rigid restricted condition, we know $\vdash_{\Box} \forall\alpha(\psi \to \Box\psi)$. So by RN, $\Box\forall\alpha(\psi \to \Box\psi)$. So by CBF, $\forall\alpha\Box(\psi \to \Box\psi)$, and by $\forall$E, $\Box(\psi \to \Box\psi)$. But (ϑ) implies $\psi \to \mathcal{A}\varphi$. From these last two results, it follows by (172.7) that $\mathcal{A}(\psi \to \varphi)$.

strict, suppose we were to apply Rule RN to conclude $\Box O!u$. Then by the conventions just described for eliminating the restricted variable, this last claim would abbreviate $O!x \to \Box O!x$. This is, in fact, a theorem (180.1), and so is its universalization $\forall x(O!x \to \Box O!x)$, by GEN. So it looks like we have reasoned validly. But although the conclusions we've reached are theorems, we've not reasoned to them by valid means. The inference from $O!u$ to $\Box O!u$ by appeal to RN is not strictly valid, for the conclusion, $\Box O!u$, i.e., $O!x \to \Box O!x$, does not validly follow by RN from the premise $O!u$, i.e., from $O!x \to O!x$. The only reason we reached a theorem, despite our invalid application of RN, is that the theorem in question, $O!x \to \Box O!x$, is derivable by other, completely valid means, as in the proof of (180.1). Analogous remarks apply to $D!x$.

To see how the application of RN to theorems with free restricted variables could turn out to be disastrous, consider an analogous piece of reasoning in connection with classes. Clearly, it is a modally strict theorem that $Class(x) \to Class(x)$. So, by the conventions described in the foregoing, we should be able to assert this as the modally strict theorem $Class(c)$. Since the theorem is modally strict, suppose we apply Rule RN to conclude that $\Box Class(c)$ is a theorem. Then by the conventions just described for eliminating the restricted variable, this last claim would abbreviate $Class(x) \to \Box Class(x)$. And since we've just apparently established it as a *theorem*, we may infer $\forall x(Class(x) \to \Box Class(x))$, by GEN. But this contradicts theorem (320.2), which explicitly asserts $\neg\forall x(Class(x) \to \Box Class(x))$. By contrast, if we had simply applied RN to the modally strict theorem with which we began, i.e., $Class(x) \to Class(x)$, then we obtain $\Box(Class(x) \to Class(x))$. And by GEN, we have $\forall x \Box(Class(x) \to Class(x))$. Both of these last conclusions are theorems and harmless.

A related problem affects Rule RA, as follows. Start again with the modally strict theorem $Class(x) \to Class(x)$, and abbreviate this as $Class(c)$. Then Rule RA implies $\mathcal{A}Class(c)$ is a theorem. So by the conventions just described for eliminating the restricted variable, this last claim would abbreviate $Class(x) \to \mathcal{A}Class(x)$. And since we've just apparently established it as a *theorem*, we may infer $\forall x(Class(x) \to \mathcal{A}Class(x))$, by GEN, and then $\Box\forall x(Class(x) \to \mathcal{A}Class(x))$, by RN. But theorem (320.3) is $\neg\Box\forall x(Class(x) \to \mathcal{A}Class(x))$.

Thus, *any* modal reasoning with Rules RN and RA has to be undertaken with great care if there are free restricted variables present; this calls into question whether we can invoke instances, with such variables, of any theorems derived with the help of RN and RA. Moreover, there are issues with other axioms and theorems as well. For example, one might wonder, when c is the only restricted variable occurring free in φ, whether the following is a legitimate instance of axiom (47) and whether it is a necessary truth:

$$c = \imath c\varphi \equiv \forall c'(\mathcal{A}\varphi_c^{c'} \equiv c' = c)$$

This is interesting because the restricted variable c has both bound and free occurrences. Let consider an example, say, when φ is $UnionOf(c,c',c'')$, i.e., c is the union of c' and c'', in (338), and consider the fact that a description such as $\imath c UnionOf(c,c',c'')$ has free restricted variables c' and c'' that are governed by restriction conditions that are not within the scope of the $\imath$-operator, whereas the bound variable c is governed by a restriction condition that is within the scope of that operator. So as example of the question above, we have to consider whether the following, with both bound and free occurrences of c, is an instance of necessary axiom (47) and whether it is subject to Rule RN:

$$c = \imath c UnionOf(c,c',c'') \equiv \forall c'''(\mathcal{A}UnionOf(c''',c',c'') \equiv c''' = c)$$

Similarly, consider theorem (176.2). If c is the only free restricted variable in φ, then it is not obvious whether the following is an instance of this theorem:

$$\imath c\varphi\downarrow \equiv \mathcal{A}\exists! c\varphi$$

And it is not obvious whether the following, with additional free restricted variables, is also an instance:

$$\imath c UnionOf(c,c',c'')\downarrow \equiv \mathcal{A}\exists! c UnionOf(c,c',c'')$$

In all these cases, the use of free restricted variables increases cognitive load; one has to spend significant cognitive resources to answer one or more of these questions and this can be distracting when one is in the middle of reasoning for some conclusion. We leave further discussion of these particular questions to an appendix to this chapter (Appendix IV). For now it should be clear that there are real concerns about reasoning with free restricted variables.

To address these concerns, we adopt two methodologies, i.e., two sets of conventions; one is for free variables introduced by non-rigid restriction conditions, while the other is for free variables introduced by rigid restriction conditions:

(.1) **Reasoning with Free Restricted Variables** (Non-rigid)
In the remainder of this text, we:

(a) officially banish free restricted variables from the statement of axioms and theorems and use only bound restricted variables when those can simplify the statement of the principle in question (informally, however, we may use *free* restricted variables only when we specify that they are to be interpreted under the *generality interpretation*, as discussed below), and

(b) as the first step of any derivation or proof, eliminate the bound restricted variables so as to avoid reasoning with premises and conclusion that contain free restricted variables; thus we never apply Rules RN and RA to formulas with free restricted variables when these are non-rigid.

Note how (.1.a) and (.1.b) combine to avoid the problems discussed above. By (a), we may not abbreviate the theorem $Class(x) \rightarrow Class(x)$ as $Class(c)$. Instead, since $\forall x(Class(x) \rightarrow Class(x))$ is a theorem that has a form that can be abbreviated with bound restricted variables, then we may either abbreviate this as $\forall c Class(c)$ or stipulate that $Class(c)$ is to be understood as $\forall c Class(c)$, thereby giving the variable c the *generality interpretation*. If we apply RN or RA to $\forall c Class(c)$, to obtain $\Box\forall c Class(c)$ or $\mathcal{A}\forall c Class(c)$, no untoward consequences arise. When expanded, these last claims assert $\Box\forall x(Class(x) \rightarrow Class(x))$ and $\mathcal{A}\forall x(Class(x) \rightarrow Class(x))$, both of which are trivially true.

To take a more interesting example, consider the Principle of Extensionality (343), which we shall prove in the next section:

$$(Class(x) \,\&\, Class(y)) \rightarrow (x = y \equiv \forall z(z \in x \equiv z \in y)) \qquad (343)$$

By GEN, it is a theorem that:

$$\forall x \forall y((Class(x) \,\&\, Class(y)) \rightarrow (x = y \equiv \forall z(z \in x \equiv z \in y))) \qquad (343)$$

This has a form that we can abbreviate with bound restricted variables, so we may either assert the above as:

(ϑ) $\forall c \forall c'(c = c' \equiv \forall z(z \in c \equiv z \in c'))$

or assert the following under the generality interpretation:

(ζ) $c = c' \equiv \forall z(z \in c \equiv z \in c')$

Under the generality interpretation of the variables, (ζ) simply abbreviates (ϑ). So by adopting either of these two expediencies, we can't get into trouble by applying RN or RA to an axiom or theorem. For if we apply RN or RA to (ϑ) and then eliminate the bound restricted variable, we obtain:

$$\Box\forall x \forall y((Class(x) \,\&\, Class(y)) \rightarrow (x = y \equiv \forall z(z \in x \equiv z \in y)))$$

$$\mathcal{A}\forall x \forall y((Class(x) \,\&\, Class(y)) \rightarrow (x = y \equiv \forall z(z \in x \equiv z \in y)))$$

These follow validly from theorem (343) by two applications of GEN and one of RN or RA.

Clearly, this first method ensures that Rules RN and RA are never improperly applied to axioms or theorems that contain free restricted variables. Moreover, part (.1.b) of this method ensures that we don't formulate instances of axioms like (47) and theorems like (176.2) unless we know that they or their closures have a form that can be abbreviated with bound restricted variables.

The second method for reasoning with free restricted variables applies to those introduced by rigid restriction conditions; these allow us to both assert axioms and theorems with such variables and to reason with them, as follows:

(.2) **Reasoning With Free Restricted Variables** (Rigid)
If γ is a rigid restricted variable introduced in connection with the *rigid* restriction condition ψ:

(.a) assert axioms and theorems containing free occurrences of γ (giving them a conditional interpretation), and

(.b) when reasoning, (i) we may at any point take $\psi(\gamma)$ as a necessary axiom, (ii) we may regard the principles of quantification theory, e.g., GEN, Existential Introduction (101), Existential Elimination (102) as valid for such variables, and (iii) we may regard the modal Rules RN and RA as valid for such variables; in cases (ii) and (iii), we give any premises a conjunctive interpretation and the conclusion a conditional interpretation.

We leave examples of (.2.b.i) and (.2.b.ii) to the reader, and conclude with an extended discussion of a fact about Rules RN and RA:

(.3) **Metatheorem:** Rules RN, RA, and Rigid Restricted Variables.
Let $\gamma_1,\ldots,\gamma_n$ be rigid restricted variables introduced by the *rigid* restriction conditions $\psi_1,\ldots,\psi_n$, respectively, where $n \geq 1$ and each ψ_i has α_i as its single, free unrestricted variable, for $1 \leq i \leq n$. Suppose some possibly empty subset of $\alpha_1,\ldots,\alpha_n$ occur free in the formulas in Γ, and a possibly distinct (also possibly empty) subset of $\alpha_1,\ldots,\alpha_n$ occur free in φ. Let $\varphi(\gamma_1,\ldots,\gamma_n)$, $\Gamma(\gamma_1,\ldots,\gamma_n)$, $\Box\Gamma(\gamma_1,\ldots,\gamma_n)$, and $\mathcal{A}\Gamma(\gamma_1,\ldots,\gamma_n)$, respectively, be the result of substituting γ_i for any free occurrences of α_i in φ and in the formulas of Γ, $\Box\Gamma$, and $\mathcal{A}\Gamma$, for $1 \leq i \leq n$.[233] Then the following metarules are justified:

(.a) If $\Gamma(\gamma_1,\ldots,\gamma_n) \vdash_\Box \varphi(\gamma_1,\ldots,\gamma_n)$, then $\Box\Gamma(\gamma_1,\ldots,\gamma_n) \vdash_\Box \Box\varphi(\gamma_1,\ldots,\gamma_n)$

(.b) If $\Gamma(\gamma_1,\ldots,\gamma_n) \vdash \varphi(\gamma_1,\ldots,\gamma_n)$, then $\mathcal{A}\Gamma(\gamma_1,\ldots,\gamma_n) \vdash \mathcal{A}\varphi(\gamma_1,\ldots,\gamma_n)$

where the premises in $\Gamma(\gamma_1,\ldots,\gamma_n)$, $\Box\Gamma(\gamma_1,\ldots,\gamma_n)$, and $\mathcal{A}\Gamma(\gamma_1,\ldots,\gamma_n)$ are given the conjunctive interpretation, and the conclusions $\varphi(\gamma_1,\ldots,\gamma_n)$, $\Box\varphi(\gamma_1,\ldots,\gamma_n)$, and $\mathcal{A}\varphi(\gamma_1,\ldots,\gamma_n)$ are given the conditional interpretation.

A few observations are in order. First, the proof given in the Appendix adopts the following strategy: we use our conventions to unpack the abbreviations in

[233] Recall that $\Box\Gamma$ is the set of necessitations of all the formulas in the set Γ, i.e., $\{\Box\chi \mid \chi \in \Gamma\}$, and $\mathcal{A}\Gamma$ is the set of actualizations of all the formulas in the set Γ, i.e., $\{\mathcal{A}\chi \mid \chi \in \Gamma\}$. So:

$\Gamma(\gamma_1,\ldots,\gamma_n)$ is the set of formulas of the form $\chi(\gamma_1,\ldots,\gamma_n)$ such that $\chi \in \Gamma$.

$\Box\Gamma(\gamma_1,\ldots,\gamma_n)$ is the set of formulas of the form $\Box\chi(\gamma_1,\ldots,\gamma_n)$ such that $\chi \in \Gamma$.

$\mathcal{A}\Gamma(\gamma_1,\ldots,\gamma_n)$ is the set of formulas of the form $\mathcal{A}\chi(\gamma_1,\ldots,\gamma_n)$ such that $\chi \in \Gamma$.

the metarule (i.e., eliminate the free restricted variables) and then show that the resulting metarule is justified. The conjunctive interpretation has to be given to each formula in each premise set.[234] But when we give the conditional interpretation to the conclusion $\varphi(\gamma_1,\ldots,\gamma_n)$, then it doesn't matter whether we abbreviate it as $(\psi_1 \mathbin{\&} \ldots \mathbin{\&} \psi_n) \to \varphi$ or as $\psi_1 \to (\ldots \to (\psi_n \to \varphi)\ldots)$. They are equivalent. Similarly, it doesn't matter whether the conclusion $\Box\varphi(\gamma_1,\ldots,\gamma_n)$ abbreviates $(\psi_1 \mathbin{\&} \ldots \mathbin{\&} \psi_n) \to \Box\varphi$ or $\psi_1 \to (\ldots \to (\psi_n \to \Box\varphi)\ldots)$. And similar considerations apply to conclusions of the form $\mathcal{A}\varphi(\gamma_1,\ldots,\gamma_n)$. Once these abbreviations are unpacked, the proof in the Appendix shows that the resulting metarule is justified.

Second, note that (.3.a), the analogue metarule for RN, is the *strong* version of the rule. By the reasoning in footnote 132, the weak version of the metarule will follow from the strong version, where the weak version of the metarule is:

If $\Gamma(\gamma_1,\ldots,\gamma_n) \vdash_\Box \varphi(\gamma_1,\ldots,\gamma_n)$, then $\Box\Gamma(\gamma_1,\ldots,\gamma_n) \vdash \Box\varphi(\gamma_1,\ldots,\gamma_n)$

The only difference here is that in the consequent of the metarule, the derivability relation is $\vdash$ rather than $\vdash_\Box$.

Third, when $\Gamma(\gamma_1,\ldots,\gamma_n)$ is empty, then (.3.a) and (.3.b) reduce, respectively, to the metarules:

If $\vdash_\Box \varphi(\gamma_1,\ldots,\gamma_n)$, then $\vdash_\Box \Box\varphi(\gamma_1,\ldots,\gamma_n)$

If $\vdash \varphi(\gamma_1,\ldots,\gamma_n)$, then $\vdash \mathcal{A}\varphi(\gamma_1,\ldots,\gamma_n)$

Here it is easy to see and state what these abbreviate:

If $\vdash_\Box (\psi_1 \mathbin{\&} \ldots \mathbin{\&} \psi_n) \to \varphi$, then $\vdash_\Box (\psi_1 \mathbin{\&} \ldots \mathbin{\&} \psi_n) \to \Box\varphi$,
where $\psi_1,\ldots,\psi_n$ are rigid restriction conditions.

If $\vdash (\psi_1 \mathbin{\&} \ldots \mathbin{\&} \psi_n) \to \varphi$, then $\vdash (\psi_1 \mathbin{\&} \ldots \mathbin{\&} \psi_n) \to \mathcal{A}\varphi$,
where $\psi_1,\ldots,\psi_n$ are rigid restriction conditions.

The proofs of (.3.a) and (.3.b) in the Appendix imply that both of these metarules are justified.

Fourth, it would make the foregoing even clearer to see (.3.a) in action. We've chosen an example using restricted variables u, v ranging over discernible objects, though there is a similar example using restricted variables ranging over ordinary objects. The following is an instance of (.3.a):

(A) If $\forall F(Fu \equiv Fv) \vdash_\Box u =_D v$, then $\Box\forall F(Fu \equiv Fv) \vdash_\Box \Box u =_D v$

[234] When we give the premises the conjunctive interpretation, then each $\chi(\gamma_1,\ldots,\gamma_n)$ in $\Gamma(\gamma_1,\ldots,\gamma_n)$ abbreviates $\psi_1 \mathbin{\&} \ldots \mathbin{\&} \psi_n \mathbin{\&} \chi$. Each $\Box\chi(\gamma_1,\ldots,\gamma_n)$ in $\Box\Gamma(\gamma_1,\ldots,\gamma_n)$ abbreviates $\psi_1 \mathbin{\&} \ldots \mathbin{\&} \psi_n \mathbin{\&} \Box\chi$. And each $\mathcal{A}\chi(\gamma_1,\ldots,\gamma_n)$ in $\mathcal{A}\Gamma(\gamma_1,\ldots,\gamma_n)$ abbreviates $\psi_1 \mathbin{\&} \ldots \mathbin{\&} \psi_n \mathbin{\&} \mathcal{A}\chi$.

By the reasoning in the proof of (.3.a) in the Appendix and the fact that the same variables occur free in the premise and conclusion of the derivations in both the antecedent and consequent, (A) abbreviates:

(B) If $D!x, D!y, \forall F(Fx \equiv Fy) \vdash_{\Box} x =_D y$, then $D!x, D!y, \Box\forall F(Fx \equiv Fy) \vdash_{\Box} \Box x =_D y$

To see that (B) is justified, assume $D!x, D!y, \forall F(Fx \equiv Fy) \vdash_{\Box} x =_D y$. Then we have to show that there is a modally strict proof of $\Box x =_D y$ from $D!x$, $D!y$, and $\Box\forall F(Fx \equiv Fy)$. So assume $D!x$, $D!y$, and $\Box\forall F(Fx \equiv Fy)$. Since $D!x$ and $D!y$ are rigid restriction conditions (273.8), we know $D!x \to \Box D!x$ and $D!y \to \Box D!y$. Hence, $\Box D!x$ and $\Box D!y$. But it follows from our initial assumption by the original Rule RN (68) that $\Box D!x, \Box D!y, \Box\forall F(Fx \equiv Fy) \vdash_{\Box} \Box x =_D y$. Since we've established all of the premises, it follows that $\Box x =_D y$. Thus, there is a modally strict proof of $\Box x =_D y$ from $D!x$, $D!y$, and $\Box\forall F(Fx \equiv Fy)$.[235]

We can repurpose this example to show how Rule RN applies to a theorem with free but rigid restricted variables. Consider the following theorem, which is easily derivable from (273.7) (we used this example in footnote 216):

(C) $(D!x \,\&\, D!z) \to (\forall F(Fx \equiv Fz) \to x = z)$

When u and v are rigid restricted variables ranging over discernible objects, then (C) can be expressed as follows:

(D) $\forall F(Fu \equiv Fv) \to u = v$

Then the version of (.3.a) in which Γ is empty implies:

(E) $\Box(\forall F(Fu \equiv Fv) \to u = v)$

where this abbreviates:

(F) $(D!x \,\&\, D!z) \to \Box(\forall F(Fx \equiv Fz) \to x = z)$

Clearly, (F), which differs from (C) only by the presence of a single $\Box$, is a theorem:

[235]The foregoing example is not optimal for two reasons:

- The formula being derived in the antecedent of (A), $u =_D v$, is a necessary truth when true (273.21), and that might vitiate the example since its necessitation can be derived by other means. But it should be remembered that we are considering a particular instance of (.3.a), to see why a meta-inference from one derivation to another is justified.
- The very same variables appear free in the premises and conclusion. But we've formulated (.3.a) in a more general way, so that the rigid restricted variables that are free in Γ are different from the rigid restricted variables free in φ (indeed, the variables may be of different types and involve different restriction conditions).

The same issues would arise if we were to replace $D!$ by $O!$ and $=_D$ by $=_E$ in the example. So the reader is encouraged to try to construct an example without these deficiencies.

Proof. Assume $D!x \& O!y$. If we apply Rule RM (or apply Rule RN and the K axiom) to theorem (242.1), which is reproduced above, then we know:

(G) $\Box(D!x \,\&\, D!z) \to \Box(\forall F(Fx \equiv Fz) \to x = z)$

But $D!x \to \Box D!x$ and $D!z \to \Box D!z$, by (273.8). Hence we know $\Box D!x \,\&\, \Box D!z$. And this implies, by (158.3), $\Box(D!x \,\&\, D!z)$. From this last result and (G), it follows that $\Box(\forall F(Fx \equiv Fz) \to x = z)$.

Thus, we've established (F) by conditional proof, and this shows that this particular application of Rule RN to theorems with free, but rigid, restricted variables is justified. We leave it to the reader to construct examples that show Rule RA can be deployed in derivations and applied to theorems when all the free restricted variables are rigid.

By contrast, since *Class*(x) is a non-rigid condition, the restricted variables $c, c', c'', \ldots$ will be governed by the conventions for reasoning with free, non-rigid restricted variables and, when we turn to the theorems about natural classes in the next section, those conventions will direct our method of proof in the Appendix. By eschewing free but non-rigid restricted variables in derivations and proofs, we avoid the burden of establishing the validity of reasoning with terms that advertise themselves one way, but which fail to behave as advertised when the modal context changes.

(**342**) **Remark:** Digression on Restricted Variables and Empty/Weak Restriction Conditions. In (336), we saw that a *weak restriction condition* is formula ψ that has a single free variable α and for which there are proofs, but not modally strict proofs, that ψ is non-empty and has existential import. And we saw that an *empty restriction condition* is a formula ψ (with a single free variable α) that has existential import but that is empty – i.e., it is not a theorem that $\exists\alpha\psi$.

It is important to mention why we shall altogether avoid introducing variables in connection with empty restriction conditions, and minimize the introduction of restricted variables in connection with weak restriction conditions. The problems for restricted variables introduced for empty restriction conditions are: (a) such variables don't obey axiom (39.2) and (b) the conditional $\forall\gamma\varphi \to \exists\gamma\varphi$ (where γ is restricted) would be false in the situation where $\neg\exists\alpha\psi$, even though it has the form of a logical theorem. To see (a), note that if γ is a variable introduced in connection with an *empty* restriction condition, then $\gamma\downarrow$ wouldn't be axiomatic. Thus we would have a class of variables that don't obey axiom (39.2). To see (b), consider claim:

(ϑ) $\forall\gamma\varphi^{\gamma}_{\alpha} \to \exists\gamma\varphi^{\gamma}_{\alpha}$

This claim has the form of a logical theorem: $\forall\beta\varphi \to \exists\beta\varphi$ (103.1). But it fails to hold in the case where in fact $\neg\exists\alpha\psi$. To see why, note that by the conventions for bound restricted variables given in (337.1), (ϑ) is shorthand for:

(ξ) $\forall\alpha(\psi \to \varphi) \to \exists\alpha(\psi \,\&\, \varphi)$

Then, under the hypothesis $\neg\exists\alpha\psi$, one can derive the antecedent of (ξ) and the negation of its consequent. From $\neg\exists\alpha\psi$, i.e., $\forall\alpha\neg\psi$, it would follow that $\neg\psi$; hence $\psi \to \varphi$, and so $\forall\alpha(\psi \to \varphi)$, by GEN. But if $\neg\exists\alpha\psi$, then $\neg\exists\alpha(\psi \,\&\, \varphi)$. Thus, we would have a class of variables for which the classical quantifier law $\forall\beta\varphi \to \exists\beta\varphi$ fails to be generally true. So if one were to introduce a restricted variable γ to range over those αs such that ψ without having a proof that $\exists\alpha\psi$, one would have to be much more careful when reasoning with them.

We're a little better off with weak restriction conditions and weak restricted variables. For such conditions ψ, it is provable, but not modally strictly provable, that $\exists\alpha\psi$. So if we introduce γ as a restricted variable ranging over the entities such that ψ, then the quantifier law $\forall\gamma\varphi \to \exists\gamma\varphi$ will be valid, but it will not be a modally strict law, since it will fail in modal contexts in which $\neg\exists\alpha\psi$. So when we introduce weak restricted variables in what follows, we should remember that when reasoning in modal contexts, the usual quantifier laws may fail. We are, of course, safe in the case of $Class(x)$, since we've established $\exists x Class(x)$ (315.4) as a modally strict theorem. So $Class(x)$ is restriction condition, though not a weak or empty one. It is a (modally strict) theorem that $\forall c\varphi^c_x \to \exists c\varphi^c_x$.[236]

10.6 The Laws of Natural Classes and Logical Sets

With the foregoing understanding of free restricted variables, we now officially make use of $c, c', c'', \ldots$ to range over classes, as the latter were defined in (312.2). Since $Class(x)$ is not a rigid restricted condition, we must assert the universal closures of any theorems asserted with variables $c, c', \ldots$ or take the generality interpretations of these variables, as discussed in (341). But since we've established that classes exist, the principle $\forall c\varphi \to \exists c\varphi$ holds, and we are assured that $c\!\downarrow$, $c'\!\downarrow$, are always axiomatic.

It should be noted that the laws of natural classes and logical sets derived below are *not* intended to capture the *iterative* conception of membership; the axiom for power sets will not be derivable. The iterative conception, as captured by Zermelo-Fraenkel set theory, is part of theoretical mathematics and, as such, its objects will be discussed in Chapter 15.

(**343**) **Theorem:** The Principle of Extensionality. For any classes c and c', c and c' are identical if and only if they have the same members:

[236] By (315.4), we know $\exists x Class(x)$. So let a be an arbitrary such individual, so that $Class(a)$. Now assume $\forall c\varphi^c_x$. Then by our convention, $\forall x(Class(x) \to \varphi)$. In particular, $Class(a) \to \varphi^a_x$. Hence, φ^a_x. Conjoining what we know, we may infer $Class(a) \,\&\, \varphi^a_x$. So by $\exists$I, $\exists x(Class(x) \,\&\, \varphi)$, which conclusion remains once we discharge our assumption that $Class(a)$ by $\exists$E. So by our convention, $\exists c\varphi^c_x$.

$$\forall c \forall c'(c = c' \equiv \forall z(z \in c \equiv z \in c'))$$

This theorem shows our definitions of natural classes (i.e., logically-conceived sets) and membership in a class preserve a fundamental and intuitive principle about the individuation of these objects.

(**344**) **Definition:** Natural Empty Classes. We say that a class c is *empty* just in case it has no members:

$$Empty(c) \ \equiv_{df} \ \neg\exists y(y \in c)$$

(**345**) **Theorems:** There Exists a (Unique) Empty Class.

(.1) $\exists c Empty(c)$

(.2) $\exists! c Empty(c)$

(.2) uses bound restricted variables in the context of a unique-existence quantifier. To interpret it, eliminate the uniqueness quantifier first and then apply the metadefinition in (337.2).[237] Though, of course, one could just as well add the convention that $\exists! c Empty(c)$ abbreviates $\exists! x(Class(x) \,\&\, Empty(x))$.

The reader should confirm why it is that one can validly apply the Rule of Actualization to (.2) and then apply (176.2) to yield the modally strict conclusion:

(.3) $\imath c Empty(c)\downarrow$

(**346**) **Definition:** Given (345.3), we may introduce the empty class symbol $\varnothing$ to designate the empty class:

$$\varnothing \ =_{df} \ \imath c Empty(c)$$

(**347**) **Theorem:** $\varnothing$ Exists and Is Canonical.

(.1) $\varnothing\downarrow$

(.2) $\varnothing = \imath x(A!x \,\&\, \forall F(xF \equiv \neg\exists z Fz))$

Though (.1) is a simple consequence of our definitions and theorems, we've taken the trouble to formulate it because of its philosophical significance. In general, the existence of abstract objects seems to confound a number of philosophers, and the empty class in particular is an egregious example; its existence offends the naturalist. But object theory offers a way to understand the

[237] To be maximally explicit, (.2) abbreviates $\exists c(Empty(c) \,\&\, \forall c'(Empty(c') \to c' = c))$. This in turn abbreviates: $\exists x(Class(x) \,\&\, Empty(x) \,\&\, \forall z((Class(z) \,\&\, Empty(z)) \to z = x))$. In other words, there is a class x that is empty and such that every class that is empty is identical to x.

empty class as a natural, (objectified) pattern of properties, namely, the pattern consisting of (i.e., encoding) all those properties that are unexemplified.[238] In object theory, (.2) provably identifies the null set with that objectified pattern.

(**348**) **Definition:** Universal* Classes. We've already defined the notion of a universal object; in (263.2) we defined: *Universal*(x) iff x is an abstract object that encodes every property. To avoid clash of notation, we introduce a different expression to define a universal class. We say that a class c is *universal** just in case every individual is a member of c:

$$Universal^*(c) \equiv_{df} \forall y(y \in c)$$

(**349**) **Theorems:** There Exists a (Unique) Universal* Class.

(.1) $\exists c Universal^*(c)$

(.2) $\exists! c Universal^*(c)$

(**350**) **Definition:** Unions. We say that a class c is a *union of* c' and c'' just in case the elements of c are precisely the elements of c' supplemented by the elements of c'':

$$UnionOf(c, c', c'') \equiv_{df} \forall y(y \in c \equiv (y \in c' \vee y \in c''))$$

Note that we explained how to eliminate the free restricted variables from this particular definition in (338.1).

(**351**) **Theorems:** Existence of (Unique) Unions. It now follows that (.1) for any classes c' and c'', there is a class c that is their union; and (.2) for any classes c' and c'', there is a unique class that is their union:[239]

(.1) $\forall c' \forall c'' \exists c UnionOf(c, c', c'')$

(.2) $\forall c' \forall c'' \exists! c UnionOf(c, c', c'')$

(**352**) $\star$**Theorem:** The Union of Two Classes Exists. It is a theorem, but not a modally strict one, that for any classes c' and c'', the class that is a union of c' and c'' exists:

[238] Such a pattern clearly exists; we encounter such properties not just in logic but also in the pursuit of science—just consider properties like *being a molecule of DNA* at a time before any such molecules existed. Surely there are properties like that now, and $\neg\exists zFz$ defines a pattern consisting of those properties.

[239] If we eliminate the unique-existence quantifier, then where x, y, z, w are all unrestricted variables, (.2) is short for:

$$\forall x \forall y(Class(x) \& Class(y) \rightarrow \exists z(Class(z) \& UnionOf(z, x, y) \& \forall w(Class(w) \& UnionOf(w, x, y) \rightarrow w = z)))$$

I.e., for any two classes x and y, there is a *unique* class that is their union.

(.1) $\forall c'\forall c''(\iota c UnionOf(c,c',c'')\downarrow)$

This can't be necessitated. For given theorem (320.3), we may intuitively reason that from the fact that x and y are classes at some other possible world, it doesn't follow that they are classes at the actual world; without such a guarantee, there is no guarantee that there is something at the actual world that is the union of x and y, since unions are classes that can only be formed from two classes (350). So from the fact that x and y are classes at some other possible world, the existence of *the* z, which (in fact) is a class and a union of x and y, is not assured.

For a purely object-theoretic account of why this theorem can't be necessitated, we simply prove the following:

(.2) $\neg\Box\forall c'\forall c''(\iota c UnionOf(c,c',c'')\downarrow)$

That is, it is not necessarily the case that for any classes c' and c'', the union of c' and c'' exists.

(**353**) **Theorem:** Modally Strict Version of the Foregoing. The previous theorem can be proved by modally strict means:

$$\neg\Box\forall c'\forall c''(\iota c UnionOf(c,c',c'')\downarrow)$$

(**354**) **Definition:** Notation for the Union of Two Classes. Given (352.1)★, we may introduce the standard notation $c' \cup c''$ to denote the union of classes c' and c'':

$$c' \cup c'' =_{df} \iota c UnionOf(c,c',c'')$$

Since we proved that $\iota c UnionOf(c,c',c'')$ exists (352.1)★ by a non-modally strict proof, the (derivation of the) identity $c' \cup c'' = \iota c UnionOf(c,c',c'')$ from the primitive Rule of Definition by Identity (73) or the derived Rule of Identity by Definition (120.1) is not modally strict. See the discussion in Remark (284). And see Remarks (339.1) and (337.3) for a full discussion of the conventions for eliminating the free and bound restricted variables in this definition. **Exercise:** Suppose we have extended our language with name s (Socrates) and that we have extended our system with the (necessary) axiom that Socrates is an ordinary object ($O!s$). Prove that $\neg(s \cup \varnothing)\downarrow$. That is, prove that the term $s \cup \varnothing$ isn't significant. [Hint: See the discussion at the end of Remark (339).]

(**355**) ★**Theorem:** The Principle of Unions. For any classes c' and c'', an object z is an element of $c' \cup c''$ if and only if z is an element of c' or an element of c'':

$$\forall c'\forall c''\forall z(z \in c' \cup c'' \equiv (z \in c' \vee z \in c''))$$

(**356**) **Definition:** Class Complements. Let us say that c' is a *class complement of* c just in case the members of c' are all and only those individuals that fail to be members of c. Then, if we write $\kappa \notin \kappa'$ instead of $\neg(\kappa \in \kappa')$, we may formalize our definition as:

$$ComplementOf(c', c) \equiv_{df} \forall y(y \in c' \equiv y \notin c)$$

It may worth mentioning that if κ_1 is a provably empty individual term, then although we can prove that for any class c, $\kappa_1 \notin c$,[240] we can't instantiate the claim $\forall y(y \in c' \equiv y \notin c)$ to obtain $\kappa_1 \in c' \equiv \kappa_1 \notin c$. So if $ComplementOf(c', c)$, one can't derive that $\kappa_1 \in c'$ from the fact that $\kappa_1 \notin c$.

(**357**) **Theorems:** Existence of (Unique) Complements. It now follows that (.1) for any class c, there is a class c' that is its complement, and (.2) for any class c, there is a unique class c' that is its complement:

(.1) $\forall c \exists c'\, ComplementOf(c', c)$

(.2) $\forall c \exists! c'\, ComplementOf(c', c)$

(**358**) $\star$**Theorems:** The Complement of a Class Exists. By reasoning analogous to that used in the proof of (352.1)$\star$, we may now infer that for every class c, the complement of c exists:

$$\forall c(\imath c' Complement(c', c)\downarrow)$$

(**359**) **Definition:** Intersections. We say that c is an *intersection* of c' and c'' just in case c has as members all and only the individuals that are common members of c' and c'':

$$IntersectionOf(c, c', c'') \equiv_{df} \forall y(y \in c \equiv y \in c' \,\&\, y \in c'')$$

(**360**) **Theorem:** Existence of (Unique) Intersections. It follows that (.1) for any classes c' and c'', there is a class c that is their intersection; and (.2) for any classes c' and c'', there is a unique class c that their intersection:

(.1) $\forall c' \forall c'' \exists c IntersectionOf(c, c', c'')$

(.2) $\forall c' \forall c'' \exists! c IntersectionOf(c, c', c'')$

These are modally strict theorems.

(**361**) $\star$**Theorem:** The Intersection of Two Classes Exists. By contrast, it is a theorem, but not a modally strict one, that for any two classes c' and c'', the intersection of c' and c'' exists:

---

[240] From $\neg\kappa_1\downarrow$, it follows by axiom (39.5.a) that $\neg G\kappa_1$. Since this holds for every G, it can then be established that $\neg\exists G(ExtensionOf(x, G) \,\&\, G\kappa_1)$. Since this holds for every x, we can infer from the definition of $\in$ (316) that $\neg\exists x(\kappa_1 \in x)$. *A fortiori*, $\neg\exists c(\kappa_1 \in c)$, for any c.

$\forall c'\forall c''(\imath c IntersectionOf(c,c',c'')\downarrow)$

Exercise: Show by either modally strict, or non-modally strict means, that it is not necessary that for any classes c' and c'', the intersection of c and c'' exists, i.e., $\neg\Box\forall c'\forall c''(\imath c IntersectionOf(c,c',c'')\downarrow)$.

(**362**) **Definition:** Notation for the Intersection of Two Classes. Given (361)★, we introduce the notation $c' \cap c''$ to denote the intersection of c' and c'':

$$c' \cap c'' =_{df} \imath c IntersectionOf(c,c',c'')$$

Note that since the proof of $(\imath c IntersectionOf(c,c',c'')\downarrow$ in (361)★ is not modally strict, the proof of the identity $c' \cap c'' = \imath c IntersectionOf(c,c',c'')$ isn't either. Moreover, given the inferential role of definitions-by-= and our conventions for restricted variables, expressions of the form $\kappa \cap \kappa'$ are significant only when we know, either by proof or by hypothesis, that $Class(\kappa)$ and $Class(\kappa')$. Thus, $\cap$ is a binary functional symbol that produces significant terms only when both arguments are significant and denote classes.

(**363**) ★**Theorem:** The Principle of Intersections. It is now straightforward to show that for any classes c' and c'', an individual z is a member of $c' \cap c''$ if and only if z is a member of both c' and c'':

$$\forall c'\forall c''\forall z(z \in c' \cap c'' \equiv (z \in c' \,\&\, z \in c''))$$

(**364**) **Theorems:** Conditional Class Comprehension. In the present system, we obtain a conditional comprehension schema that is immune to Russell's paradox if we limit the schema to those matrices that are guaranteed to define properties. Thus, (.1) if *being a y such that* φ exists, then there is a class whose members consist of all and only the individuals such that φ; (.2) if *being an y such that* φ exists, then there is a unique class whose members consist of all and only the individuals such that φ; and (.3) if *being an y such that* φ exists, then the class of individuals such that φ exists:

(.1) $[\lambda y\, \varphi]\downarrow \rightarrow \exists c\forall y(y \in c \equiv \varphi)$, provided φ has no free occurrences of c

(.2) $[\lambda y\, \varphi]\downarrow \rightarrow \exists! c\forall y(y \in c \equiv \varphi)$, provided φ has no free occurrences of c

(.3) $[\lambda y\, \varphi]\downarrow \rightarrow \imath c\forall y(y \in c \equiv \varphi)\downarrow$, provided φ has no free occurrences of c

By convention, (.1) is shorthand for $[\lambda y\, \varphi]\downarrow \rightarrow \exists z(Class(z) \,\&\, \forall y(y \in z \equiv \varphi))$, provided φ has no free occurrences of z. And analogously for (.2) and (.3). It is also worth noting that by (63.10), it follows from the fact that (.3) is a theorem that $[\lambda y\, \varphi]\downarrow \vdash \imath c\forall y(y \in c \equiv \varphi)\downarrow$.

(**365**) **Metadefinitions:** ν-Predicable Formulas and the Restricted Metavariables That Range Over Them. The foregoing group of theorems exhibits a certain pattern that we can exploit. Where ν is any individual variable, let us say:

- A formula φ is *ν-predicable* if and only if there is a proof that *being an individual ν such that φ* exists, i.e., if and only if $\vdash [\lambda\nu\,\varphi]\downarrow$.

For present purposes, we need not generalize the above definition to define $\nu_1,\ldots,\nu_n$-predicable formulas.

Clearly, there are ν-predicable formulas — any unary λ-expression of the form $[\lambda\nu\,\varphi]$ asserted to exist by (39.2) has a ν-predicable matrix. This includes those λ-expressions of the form $[\lambda\nu\,\varphi]$ in which φ taken by itself contains no free occurrences of ν. Another source of ν-predicable formulas arises when we can prove $[\lambda\nu\varphi]\downarrow$ by means of axiom (49) or by means of the Kirchner Theorem (271.1). Just as clearly, there are formulas that fail to be ν-predicable for some variable ν. If φ is a formula such as the matrix in the Clark/Boolos paradox, namely $\exists G(xG \,\&\, \neg Gx)$, then φ is not x-predicable. And if φ is the formula involved in the McMichael/Boolos paradox, namely $x = y$, then φ is neither x- nor y-predicable.

In what follows, let us adopt the conventions that:

- $\varphi[\nu], \psi[\nu], \ldots$ are metavariables that range over ν-predicable formulas.

- The use of $\varphi[\nu]$ to express a theorem schema serves to indicate that the theorem is conditional on $[\lambda\nu\,\varphi]\downarrow$.

For example, $\vdash \ldots\varphi[y]\ldots$ is shorthand for $\vdash [\lambda y\,\varphi]\downarrow \rightarrow \ldots\varphi\ldots$.

Thus, we may rewrite the theorems in (364) about Conditional Class Comprehension as follows. (.1) there is a class whose members consist of all and only the individuals y such that $\varphi[y]$; (.2) there is a unique class whose members consist of all and only the individuals y such that $\varphi[y]$; and (.3) the class of individuals y such that $\varphi[y]$ exists:

(.1) $\exists c\forall y(y \in c \equiv \varphi[y])$, provided $\varphi[y]$ has no free occurrences of c

(.2) $\exists! c\forall y(y \in c \equiv \varphi[y])$, provided $\varphi[y]$ has no free occurrences of c

(.3) $\imath c\forall y(y \in c \equiv \varphi[y])\downarrow$, provided $\varphi[y]$ has no free occurrences of c

(**366**) **Definition:** Class Abstracts. By (364.3) and our conventions in (365), it is a theorem that $\imath c\forall y(y \in c \equiv \varphi[y])\downarrow$, for any y-predicable formula $\varphi[y]$ in which c doesn't occur free. So we could introduce the class abstract notation $\{y \,|\, \varphi[y]\}$ ("the class of individuals y such that $\varphi[y]$") and be assured that it is significant. However, our theory of definitions-by-= allows us to give the following, more general definition of class abstracts, where φ is any formula:

$\{y \,|\, \varphi\} \;=_{df}\; \imath c\forall y(y \in c \equiv \varphi)$, where φ has no free occurrences of c

One of the distinctive features of this definition is that there are instances of the definiens that fail to be significant, and so the corresponding instance of the definiendum fails to be significant. See the discussion in Remark (367).

Observe also that in the above definition, φ may have free variables other than y. If φ is y-predicable and y is the only variable that has a free occurrence in φ, then $\{y|\varphi\}$ is a closed term that denotes a particular class if $\imath c\forall y(y\in c \equiv \varphi)$ is significant. But if φ is y-predicable and there are variables other than y with free occurrences in φ, then $\{y|\varphi\}$ will denote a class relative to the values of the other free variables whenever $\imath c\forall y(y\in c \equiv \varphi)$ denotes a class relative to those values of the free variables.

(**367**) **Remark:** Defined Variable-Binding Term-Forming Operators. There are two distinctive features of definition (366). The first is that the expression being introduced, $\{y\,|\ldots\}$ doesn't take variables as arguments but rather binds the occurrences of those variables in any formula φ that appears within the ellipsis (i.e., within its scope). (We haven't discussed this kind of term previously, for the term-forming operators defined heretofore haven't been variable-binding operators, nor have they had formulas within their scope; instead they have had a free variable as an *argument*.) The first occurrence of the variable y in $\{y\,|\,\varphi\}$ is not a free occurrence, but rather part of the variable-binding operator and serves to indicate which of the variables with free occurrences in φ are bound by the operator (just as the first occurrence of x in the expression $[\lambda x\,\varphi]$ is not a free occurrence but rather bound by the λ and serves to indicate which of the variables with free occurrences in φ are bound by the λ). Thus, $\{y\,|\,\varphi\}$ is rather different from such terms as $\overline{F}$, $\circ p$, ϵG, $c' \cup c''$, $c' \cap c''$, etc.

The second distinctive feature of (366) is that the definiens, $\imath c\forall y(y \in c \equiv \varphi)$, is a provably empty term for some formulas φ. For example, the definite description $\imath c\forall y(y \in c \equiv \exists G(yG \,\&\, \neg Gy))$ doesn't denote anything, not even the empty set.[241] Consequently, the second conjunct of the Rule of Definition by Identity (73) implies that $\neg\{y\,|\exists G(yG \,\&\, \neg Gy)\}{\downarrow}$. This is to be expected since the

[241] To show that $\imath c\forall y(y \in c \equiv \exists G(yG \,\&\, \neg Gy))$ is a provably empty term, it suffices to show:

$$\neg\exists x(Class(x) \,\&\, \forall y(y \in x \equiv \exists G(yG \,\&\, \neg Gy)))$$

For reductio, suppose otherwise and let a be such an object, so that we know $Class(a)$ and:

(ϑ) $\forall y(y \in a \equiv \exists G(yG \,\&\, \neg Gy))$

Since a is a class, we know by definitions (312.2) that $\exists G(ExtensionOf(a,G))$. Suppose P is such a property, so that we know $ExtensionOf(a,P)$. Then by (317.1), it follows that:

(ξ) $\forall y(y \in a \equiv Py)$

From (ϑ) and (ξ) it follows that:

(ζ) $\forall y(Py \equiv \exists G(yG \,\&\, \neg Gy))$

But now, one can use (ζ) to derive a contradiction, as in the Clark/Boolos paradox, by considering the abstract object, say b, that encodes just the property P and asking the question: does b exemplify P? It does if and only if it does not.

Clark/Boolos paradox shows that $\exists G(yG \,\&\, \neg Gy)$ is not y-predicable.

However, as long as we use a y-predicable formula $\varphi[y]$ to introduce class abstracts of the form $\{y \,|\, \varphi[y]\}$, then since $\imath c\forall y(y \in c \equiv \varphi[y])\downarrow$ is a theorem, by (364.3) and our conventions in (365), we can invoke the first conjunct of the Rule of Definition by Identity (73) and validly assert the identity $\{y \,|\, \varphi[y]\} = \imath c\forall y(y \in c \equiv \varphi[y])$ as a theorem.

(**368**) ⋆**Theorems:** Class Abstraction Principles. The classical abstraction principles now govern class abstracts built with ν-predicable formulas. (.1) y is a member of $\{y \,|\, \varphi[y]\}$ if and only if $\varphi[y]$; and (.2) z is an element of the class of individuals such that $\varphi[y]$ just in case z is such that $\varphi[y]$:

(.1) $y \in \{y \,|\, \varphi[y]\} \equiv \varphi[y]$

(.2) $z \in \{y \,|\, \varphi[y]\} \equiv \varphi[y]^z_y$, provided z is substitutable for y in $\varphi[y]$

(.2) is an analogue of a definition in *Principia Mathematica*. In Whitehead & Russell 1910–1913 [1925–1927], ∗20·02, we find $x \,\epsilon\, (\varphi\,!\,\hat{z})\,.=.\,\varphi\,!\,x$ stipulated as a definition, and we find $\vdash\!: x \,\epsilon\, \hat{z}(\psi z)\,.\equiv.\,\psi x$ asserted as a theorem (∗20·3). They read the theorem in the 2nd edition as "x is a member of the class determined by ψ when, and only when, x satisfies ψ" (1910–1913 [1925–1927, 193]).

(**369**) **Theorems:** Modally Strict Identities. We now establish a modally strict proof of (.1) the class of individuals y such that $\varphi[y]$ is identical to the extension of the property $[\lambda y\, \varphi[y]]$. From this, it follows that (.2) the class of Gs is identical to the extension of the property G:

(.1) $\{y \,|\, \varphi[y]\} = \epsilon[\lambda y\, \varphi[y]]$

(.2) $\{y|Gy\} = \epsilon G$

By the symmetry of identity, (.2) yields $\epsilon G = \{y|Gy\}$. This is another object-theoretic analogue of Carnap's semantic assertion in 1947 (19, 4-14) that "the extension of a predicator ... is the corresponding class." This latter claim was part of Carnap's extended definition of the general semantic notion of an extension, with the notion of a *class* taken as a primitive axiomatized by the mathematical theory of sets or classes. By contrast, our theorem is in the object language, with *ClassOf* defined in terms of our mathematics-free primitives.

Of course, given that we defined *ClassOf*(x, G) as *ExtensionOf*(x, G), it is to be expected that $\epsilon G = \{y|Gy\}$ is a theorem. Nevertheless, the result serves as a sanity check, given the many subsequent definitions and theorems needed to state and prove it.

(**370**) **Theorems:** The Separation Schema. Let $\varphi[y]$ be any y-predicable formula (365). Then (.1) for any class c', there is a class whose elements y are precisely the members of c' and such that $\varphi[y]$; and (.2) for any class c', there is a unique class whose elements are precisely the members of c' such that $\varphi[y]$:

(.1) $\forall c'\exists c\forall y(y\in c \equiv y\in c' \,\&\, \varphi[y])$, provided $\varphi[y]$ has no free occurrences of c

(.2) $\forall c'\exists!c\forall y(y\in c \equiv y\in c' \,\&\, \varphi[y])$, provided $\varphi[y]$ has no free occurrences of c

(**371**) ⋆**Theorems:** The Separation Set. Where $\varphi[y]$ is any y-predicable formula (365), it follows, as a non-modally strict theorem, that for any class c', the class, whose elements are precisely the members of c' such that $\varphi[y]$, exists:

$$\forall c'(\imath c\forall y(y\in c \equiv y\in c' \,\&\, \varphi[y])\downarrow), \text{ provided } \varphi[y] \text{ has no free occurrences of } c$$

(**372**) **Definition:** Class Separation Abstracts. In light of (371)⋆, the following definition:

$$\{y \,|\, y\in c' \,\&\, \varphi\} =_{df} \imath c\forall y(y\in c \equiv y\in c' \,\&\, \varphi), \text{ provided } c \text{ isn't free in } \varphi$$

yields a significant class abstract whenever φ is a y-predicable formula.

However, it is important to recall here the discussion in (120.1) of the Rule of Identity by Definition. Ignoring derivations and the premise set Γ, this rule says that if it is a (modally-strict) theorem that the definiens is significant, then the identity introduced by the definition is a (modally-strict) theorem. This is relevant here since theorem (371)⋆ tells us, with respect to the above definition, that the significance of the definiens is not a modally strict theorem. So the identity claim derived from this definition is a theorem, but not a modally strict one. Note also that this definition also introduces a variable-binding, term-forming operator, of the kind discussed in Remark (367).

(**373**) ⋆**Theorem:** Separation Abstraction Principle. Let $\varphi[y]$ be any y-predicable formula. Then the class abstracts introduced in (372)⋆ are subject to the principle that for any class c and for any object z, z is an element of the class of individuals y such that y is an element of c and such that $\varphi[y]$ if and only if z is an element of c and such that $\varphi[y]$:

$$\forall c\forall z(z\in\{y \,|\, y\in c \,\&\, \varphi[y]\} \equiv z\in c \,\&\, \varphi[y]^z_y),$$
provided z is substitutable for y in $\varphi[y]$

(**374**) ⋆**Theorem:** Consequence of Class Separation and Intersection. Where $\varphi[y]$ is any y-predicable formula, it follows that for every class c, the class whose elements are members of c and such that $\varphi[y]$ is identical to the intersection of c and the class whose elements are such that $\varphi[y]$:

$$\forall c(\{y \,|\, y\in c \,\&\, \varphi[y]\} = c\cap\{y \,|\, \varphi[y]\})$$

(**375**) **Theorem:** Collection. For any binary relation R and class c, there is a class c' whose members are precisely those objects y such that some member of c bears R to y:

$$\forall R \forall c' \exists c \forall y(y \in c \equiv \exists z(z \in c' \,\&\, Rzy))$$

Clearly, c need not be a subset of c', since the members of c' may be R-related to non-members of c'.

(**376**) **Theorem:** Not Every Object Has a Singleton. It is provable that:

$$\neg \forall x \exists c \forall y(y \in c \equiv y = x)$$

Intuitively, the formula $y = x$ is not y-predicable, on pain of contradiction; if it were y-predicable, i.e., it if were a theorem that $[\lambda y\, y = x]\downarrow$, then it would follow that $\forall x([\lambda y\, y = x]\downarrow)$, which contradicts (192.3), thereby reintroducing the McMichael/Boolos paradox. So we cannot use $y = x$ in an instance of class comprehension (364.1). Moreover, since the expression $[\lambda y\, y = x]$ doesn't denote a property, it can't be substituted for F to produce an instance of the Fundamental Theorem for Natural Classes (318). Indeed, none of the other existence principles we've derived for natural classes allow us to assert the existence of $\{y|y = x\}$. This makes sense, given the existence of indiscernible abstract objects. Nevertheless, our theory does assert that there are well-behaved singletons of discernible objects.

(**377**) **Theorems:** Natural Singletons for Discernibles. Since $D!y \,\&\, y = x$ is a y-predicable formula (273.13), the theory guarantees that (.1) there is a class whose members are the discernible objects identical to x; (.2) there is a unique class whose members are the discernible objects identical to x; and (.3) the class whose members are discernible individuals identical to x exists:

(.1) $\exists c \forall y(y \in c \equiv D!y \,\&\, y = x)$

(.2) $\exists! c \forall y(y \in c \equiv D!y \,\&\, y = x)$

(.3) $\imath c \forall y(y \in c \equiv D!y \,\&\, y = x)\downarrow$

Hence, it follows that (.4) the class of discernible individuals identical to x exists:

(.4) $\{y \mid D!y \,\&\, y = x\}\downarrow$

(**378**) **Definition:** The Natural Singleton of an Individual. In virtue of (377.4), we may define the *natural singleton*, or *natural unit class*, of x, written $\{x\}_D$, to be the class of discernible y identical to x:

$$\{x\}_D =_{df} \{y \mid D!y \,\&\, y = x\}$$

Note here that we've introduced a new functional term by way of a defined class abstraction term.

(**379**) **Theorems:** Modally Strict Facts About Singletons. It follows that (.1) the natural singleton of an indiscernible abstract object is the empty class:

(.1) $(A!z \,\&\, \neg D!z) \rightarrow \{z\}_D = \varnothing$

Consequently, if x and y are distinct but indiscernible abstract individuals, then both $\{x\}_D$ and $\{y\}_D$ are identical to the empty class and so identical. Thus, the principle, $x \neq z \rightarrow \{x\}_D \neq \{z\}_D$, fails to hold generally:

(.2) $\neg\forall x \forall z(x \neq z \rightarrow \{x\}_D \neq \{z\}_D)$

However, (.3) the natural singletons of discernible objects aren't identical with the empty class:

(.3) $D!x \rightarrow (\{x\}_D \neq \varnothing)$

Moreover, the following principle is provable, namely, that (.4) distinct discernible objects have distinct natural singletons:

(.4) $(D!x \,\&\, D!z) \rightarrow (x \neq z \rightarrow \{x\}_D \neq \{z\}_D)$

The above results show that when κ signifies a discernible object, the term $\{\kappa\}_D$ behaves as expected.[242]

(**380**) ⋆**Theorem:** Fact About Singletons. Finally, we show that a discernible object is the sole member of its natural singleton:

$$D!x \rightarrow \forall y(y \in \{x\}_D \equiv y = x)$$

(**381**) **Theorems:** Existence of Natural Pair Classes. We may prove the existence of natural pair classes; for any objects x and z, there is a (unique) class whose members are all and only those discernible individuals y that are either identical to x or identical to z:

(.1) $\exists c \forall y(y \in c \equiv D!y \,\&\, (y = x \vee y = z))$

[242] Thus, in a system like the present one, with no primitive notions of *set* or *membership* and no axioms asserting the existence of sets, the *reconstruction* of natural classes (logical sets) implies that natural singletons are well-behaved only on the discernible objects. This preserves an assumption that is part of the classical understanding of singletons; the notion of a singleton only makes sense with respect to discernibles.

It should also be mentioned that the present theory offers a second analysis of singletons when these are part of some theoretical mathematics (i.e., when asserted to exist by an axiomatic set theory). Object theory would not identify those objects along the lines indicated above. Instead, it would identify such singletons as abstract objects whose encoded properties are the ones attributed to them in the set theory in question. Such an identification would be part of the analysis of *theoretical* mathematics in Chapter 15. On that analysis, each distinct axiom system for characterizing a primitive notion of set membership (i.e., for each system that leads to a distinct body of theorems) yields a different (abstract) *membership* relation. If a set theory T attributes classical properties to singleton sets, stated in terms of the relation of identity, $=_T$, that is assumed in T, then object theory will identify the singletons described by T as abstract objects that encode their mathematical properties and so the singletons described by T encode those classical properties.

(.2) $\exists! c \forall y(y \in c \equiv D!y \,\&\, (y = x \vee y = z)$

Exercise. Show:

$$\iota c \forall y(y \in c \equiv D!y \,\&\, (y = x \vee y = z))\downarrow$$

Then define:

$$\{x, z\}_D =_{df} \iota c \forall y(y \in c \equiv D!y \,\&\, (y = x \vee y = z))$$

Then show that if x and z are discernible objects, the natural pair class $\{x, z\}_D$ behaves classically (i.e., that its sole members are x and z). That is, prove: $(D!x \,\&\, D!z) \rightarrow \forall y(y \in \{x, z\}_D \equiv y = x \vee y = z)$.

(**382**) **Theorems:** Natural Adjunction. It is a theorem that there is a (unique) class c whose elements are all and only those objects that are either the members of a given class c' or identical$_D$ to some given object x:

(.1) $\forall c' \exists c \forall y(y \in c \equiv y \in c' \vee y =_D x)$

(.2) $\forall c' \exists! c \forall y(y \in c \equiv y \in c' \vee y =_D x)$

Exercise. Find and prove a connection between $\{y \mid y \in c' \vee y =_D x\}$ and c' when x is an indiscernible abstract object. Show that $\{y \mid y \in c' \vee y =_D x\}$ behaves classically when x is ordinary.

(**383**) **Exercises:** Anti-Extensions. Let us say that x is an *anti-extension of* property G if and only if x is an abstract object that encodes exactly the properties F that are exemplified by all and only those objects that fail to exemplify G. Prove that for any G, there is a unique anti-extension of G, and define $\overline{\epsilon}G$ as *the* anti-extension of G. Show $\overline{\epsilon}G = \epsilon\overline{G}$. Formulate and prove some interesting theorems about anti-extensions.

10.7 Abstraction via Equivalence Conditions

(**384**) **Metadefinition:** Equivalence Conditions on Relations. We now observe a general pattern that has emerged in the previous two sections. Let φ be any formula in which there are free of occurrences of the *two* distinct n-place relation variables (for some n). Suppose we've distinguished these free variables from any other free variables that may occur in φ, and that we may refer to one of these distinguised variables as 'the first' if it has the first free occurrence in φ (and refer to the other as 'the second'). Then where α and β are *any* two n-ary relation variables, let us write $\varphi(\alpha, \beta)$ for the result of simultaneously substituting α for all the free occurrences of the first distinguished free variable in φ and substituting β for all the free occurrences of the second distinguished free variable in φ. Thus, if φ happens to have α and β as the two distinguished free variables, then $\varphi(\alpha, \beta)$ just is φ. Given this notational convention, we say:

Equivalence Condition:
A formula φ with two distinct n-place relation variables is an *equivalence condition* on n-ary relations whenever the following are all provable:

$$\varphi(\alpha,\alpha) \qquad \text{(Reflexivity)}$$

$$\varphi(\alpha,\beta) \to \varphi(\beta,\alpha) \qquad \text{(Symmetry)}$$

$$\varphi(\alpha,\beta) \to (\varphi(\beta,\gamma) \to \varphi(\alpha,\gamma)) \qquad \text{(Transitivity)}$$

For example, it is easy to show that the formulas $q \equiv p$ is an equivalence condition on propositions, and that $\forall z(Fz \equiv Gz)$ is an equivalence condition on properties, i.e., to prove:

- $q \equiv q$
  $(q \equiv p) \to (p \equiv q)$
  $((q \equiv p)\,\&\,(p \equiv r)) \to q \equiv r$

- $\forall z(Fz \equiv Fz)$
  $\forall z(Fz \equiv Gz) \to \forall z(Gz \equiv Fz)$
  $(\forall z(Fz \equiv Gz)\,\&\,\forall z(Gz \equiv Hz)) \to \forall z(Fz \equiv Hz)$

Intuitively speaking, an equivalence condition with free n-ary relation variables α,β partitions the domain of n-ary relations over which α,β range in much the same way that an equivalence relation on individuals partitions the domain of individuals. The formula $q \equiv p$ partitions the domain of propositions into mutually exclusive and jointly exhaustive cells of materially equivalent propositions. The formula $\forall z(Fz \equiv Gz)$ partitions the domain of properties into mutually exclusive and jointly exhaustive cells of materially equivalent properties.

When the free variables in φ are both proposition or property variables, then we can easily build objects that represent the equivalence classes of the partition.[243] Truth-values and extensions of properties (i.e., classes), as defined earlier, are both based on instances of comprehension for abstract objects involving equivalence conditions:

- We used the equivalence condition $q \equiv p$ on propositions in (286) to define *TruthValueOf*(x,p).

- We used the equivalence condition $\forall z(Fz \equiv Gz)$ on properties in (312.1) to define *ExtensionOf*(x,G).

These definitions were in turn used to define the canonical objects $\circ p$ and ϵG, respectively. This process can be generalized, as the next series of definitions and theorems show.

[243] To see the beginnings of a technique for representing the equivalence classes of relations, see Linsky &: Zalta 2006, 83, n. 27.

(**385**) **Definitions:** Abstractions from Equivalence Conditions. For the next sequence of definitions and theorems we adopt the following conventions:

- Let φ be any equivalence condition on propositions.
- Let ψ be any equivalence condition on properties.

Then we can define (.1) x is the *φ-abstraction of p* if and only if x is an abstract object that encodes all and only the properties $[\lambda y\, q]$ such that q is some proposition φ-equivalent to p; and (.2) x is the *ψ-abstraction of G* if and only if x is an abstract object, G exists, and x encodes exactly the properties that are ψ-equivalent to G:

(.1) $\varphi\text{-}AbstractionOf(x,p) \equiv_{df} A!x \,\&\, \forall F(xF \equiv \exists q(\varphi(q,p) \,\&\, F = [\lambda y\, q]))$

(.2) $\psi\text{-}AbstractionOf(x,G) \equiv_{df} A!x \,\&\, G\downarrow \,\&\, \forall F(xF \equiv \psi(F,G))$

(**386**) **Theorems:** Existence of Unique Abstractions. Where φ and ψ are equivalence conditions of the kind just defined, the Strengthened Comprehension Principle for Abstract Objects (250) yields the existence of a unique abstraction:

(.1) $\exists!x(\varphi\text{-}AbstractionOf(x,p))$

(.2) $\exists!x(\psi\text{-}AbstractionOf(x,G))$

So by (252), we can prove that the corresponding canonical descriptions are significant:

(.3) $\imath x(\varphi\text{-}AbstractionOf(x,p))\downarrow$

(.4) $\imath x(\psi\text{-}AbstractionOf(x,G))\downarrow$

(**387**) **Theorems:** Principles Underlying the Fregean Biconditionals. (.1) if x is a φ-abstraction of p and y is a φ-abstraction of q, then x is identical to y if and only if p and q satisfy φ; (.2) if x is a ψ-abstraction of G and y is a ψ-abstraction of H, then x is identical to y if and only if G and H satisfy ψ:

(.1) $(\varphi\text{-}AbstractionOf(x,p) \,\&\, \varphi\text{-}AbstractionOf(y,q)) \to (x = y \equiv \varphi(p,q))$

(.2) $(\psi\text{-}AbstractionOf(x,G) \,\&\, \psi\text{-}AbstractionOf(y,H)) \to (x = y \equiv \psi(G,H))$

These modally strict principles underlie (389) proved below.

(**388**) **Definitions:** Notation for Abstractions of Equivalence Conditions. Where φ and ψ are equivalence conditions on propositions and properties, respectively, we may introduce notation for the φ-abstraction of p and the ψ-abstraction of G:

(.1) $\widehat{p}_\varphi =_{df} \imath x(\varphi\text{-}AbstractionOf(x,p))$

(.2) $\widehat{G}_\psi =_{df} \imath x(\psi\text{-}AbstractionOf(x,G))$

Intuitively, for any equivalence condition on propositions or properties, our theory distinguishes a canonical logical object for each cell of the partition it induces. For any proposition p, $\widehat{p}_\varphi$ encodes all and only the propositions q such that $\varphi(q,p)$ and, for any property G, $\widehat{G}_\psi$ encodes all and only the properties F such that $\psi(F,G)$.

(**389**) ⋆**Theorems:** Frege's Principle for Abstractions. Our results now justify one classical form of *definition by abstraction*, since the above definitions and theorems yield the following non-modally strict theorems, where φ and ψ are equivalence conditions on propositions and properties, respectively:

(.1) $\widehat{p}_\varphi = \widehat{q}_\varphi \equiv \varphi(p,q)$

(.2) $\widehat{F}_\psi = \widehat{G}_\psi \equiv \psi(F,G)$

Finally, note that Frege's 'Julius Caesar problem' doesn't arise for any abstractions defined along the above lines. The conditions $\widehat{p}_\varphi = x$ and $\widehat{F}_\psi = x$ are defined for every x. These are simply instances of definition (23.1). We've not introduced $\widehat{p}_\varphi$ and $\widehat{F}_\psi$ as primitive notions and stipulated (.1)⋆ and (.2)⋆ as axioms. Rather, we've defined $\widehat{p}_\varphi$ and $\widehat{F}_\psi$ in terms of canonical descriptions. The identity conditions of $\widehat{p}_\varphi$ and $\widehat{F}_\psi$, with respect to any other object x, have been defined independently.

10.8 Abstraction via Equivalence Relations

We now turn from abstractions over equivalence conditions on relations to abstractions over equivalence relations on individuals.

10.8.1 Directions and Shapes

We begin by analyzing directions and shapes as natural mathematical objects (in a later chapter, when we analyze mathematical theories generally, the *theoretical* mathematical notions of *direction* and *shape* will thereby receive an analysis; but we're putting that aside here). We shall be assuming (a) an ordinary, pretheoretic understanding of the properties *being a line* and *being a figure*, and (b) an ordinary, pretheoretic understanding of the relations *being parallel to* and *being similar to*. By starting with our pretheoretic understanding of these notions, we intend to apply them to (possibly) concrete objects and not to the idealizations of geometry. Thus, we don't assume any theoretical axioms of geometrical theories in what follows.[244]

(**390**) **Remark:** Pretheoretic Conception of Lines. Consider the property *being a line* familiar to us from ordinary language. In what follows, we use L to denote this property (note that, up to now, we have used L for a different purpose, namely, as an abbreviation for the property $[\lambda x\, E!x \rightarrow E!x]$). We won't make use of any assumptions governing this property other than the modally fragile claim that lines exist ($\exists x Lx$) and they obey the assumptions governing the relation of *being parallel to* described below.[245]

Now the ordinary binary relation *being parallel to* (written '$\|$' using infix notation), is an equivalence relation restricted to individuals that exemplify the property *being a line*. That is, we assume the following principles governing *being a line* and *being parallel to*:

$$Lx \rightarrow x \| x$$

$$Lx \,\&\, Ly \rightarrow (x \| y \rightarrow y \| x)$$

$$Lx \,\&\, Ly \,\&\, Lz \rightarrow (x \| y \,\&\, y \| z \rightarrow x \| z)$$

Now since we're taking the claim that lines exist as a modally fragile assumption, let's suppose for the purposes of this section, that in fact it is a modally fragile axiom that $\exists x Lx$. Then Lx becomes a *weak* restriction condition, as defined in (336), since it is a condition with a single free variable that is provably

[244] We trust one can distinguish the pretheoretic properties and relations such as *being a line*, *being a figure*, *being parallel to*, *being similar to*, etc., from their theoretical counterparts. Whereas the theoretical counterparts are governed by the axioms of some implicit or explicit mathematical theory, the pretheoretic properties and relations are not. For now, we shall discuss only the pretheoretic properties and relations, as referenced in ordinary, everyday language. We make this a bit more explicit below.

By contrast, the theoretical counterparts will be subject to our analysis of the relations of *theoretical* mathematics in Chapter 15. In that later chapter, we assert the existence of abstract relations (including abstract properties) in addition to abstract individuals. Abstract relations will be distinguished from ordinary relations; the former, but not the latter, may encode properties of relations. Similarly, abstract properties will be distinguished from ordinary properties; the former may encode properties of properties. In Chapter 15, theoretical mathematical properties axiomatized in some mathematical theory, such as *being a number*, *being a set*, and *being a line*, will be identified as particular abstract properties (relative to the theory in question).

[245] It does no harm to fill in the pretheoretic notion of a line a bit with the following observations, since they won't play a role in what follows. But, pretheoretically, lines (e.g., lines on a piece of paper, lines in the sand, etc.) are concrete objects (i.e., $\forall x(Lx \rightarrow E!x)$), and may consist of discrete parts (e.g., a line of people) or may be continuous at a certain level of granularity (e.g., a line of ink on paper). Moreover, lines have some (approximate) physical length and (average) thickness. Of course, if the reader finds any of this controversial, she is entitled to ignore it; none of these observations are needed for understanding the analysis that follows. But I take it that Frege's discussion of *the Earth's axis* as an example of a line is an abstraction and would be analyzed as a theoretical entity of science, and hence as an abstract object, not an ordinary one.

(but not strictly provably) non-empty and provably (but not strictly provably) has existential import. It also follows that Lx is not a rigid restriction condition, as defined in (340).[246] So, over the course of the next few items, we may use the variables u, v as *weak* and *non-rigid* restricted variables ranging over lines (we continue to use the variables x, y, z as variables for any kind of object).

Hence, the principles displayed immediately above may be written as follows:

$$\forall u(u \| u)$$

$$\forall u \forall v(u \| v \to v \| u)$$

$$\forall u \forall u' \forall u''(u \| u' \,\&\, u' \| u'' \to u \| u'')$$

So let us extend object theory for the moment with these assumptions. It should be remembered here that care must be taken when reasoning with weak restricted variables in modal contexts, for the reasons mentioned in (340) and (342).

(**391**) **Lemma:** Fact About *Being Parallel To*. From the assumption that *being parallel to* is an equivalence relation on lines, we can establish that for any two lines u and u', the properties *being a line parallel to* u and *being a line parallel to* u' are materially equivalent if and only if u is parallel to u':

$$\forall u \forall u'(\forall z([\lambda v\, v \| u]z \equiv [\lambda v\, v \| u']z) \equiv u \| u')$$

This fact plays a key role in what follows.

(**392**) **Definition:** Directions of Lines. We may define: x is a *direction of* line u just in case x is an extension of the property *being a line* v *parallel to* u:

$$DirectionOf(x, u) \ \equiv_{df} \ ExtensionOf(x, [\lambda v\, v \| u])$$

By our conventions in (338.1) and (337.4), this abbreviates:

$$DirectionOf(x, y) \ \equiv_{df} \ Ly \,\&\, ExtensionOf(x, [\lambda z\, Lz \,\&\, z \| y])$$

(**393**) **Theorems:** The Conditional Existence of *Directions Of*. It now follows that (.1) every line has a direction; and (.2) every line has a unique direction:

(.1) $\forall u \exists x DirectionOf(x, u)$

[246] This is a consequence of taking $\exists xLx$ as a modally fragile axiom, since such an axiom allows for the possibility that there are no concrete lines ($\Diamond\neg\exists xLx$). For suppose that $\Diamond\neg\exists xLx$, i.e., $\neg\Box\exists xLx$. Then to see that it follows that $\neg\forall x(Lx \to \Box Lx)$, suppose otherwise, i.e., $\forall x(Lx \to \Box Lx)$, for reductio. Given our modally fragile axiom $\exists xLx$, suppose a is such an object, so that we know La. Then by our reductio hypothesis, $\Box La$. Hence $\exists x\Box Lx$, and so by the Buridan formula (168.1), $\Box\exists xLx$, which contradicts the assumption $\Diamond\neg\exists xLx$. So we can't regard Lx as a rigid restriction condition, since the definition in (340) requires that there be a (modally strict) proof that $\forall x(Lx \to \Box Lx)$.

(.2) $\forall u \exists! x DirectionOf(x,u)$

Note that (.1) and (.2) abbreviate quantified conditionals, since the variable u is restricted.

(**394**) **Theorem:** Principle Underlying the Fregean Abstraction Principle for Directions. If x is a direction of line u and y is a direction of line v, then x is identical to y if and only if u is parallel to v:

$$(DirectionOf(x,u) \,\&\, DirectionOf(y,v)) \rightarrow (x=y \equiv u \parallel v)$$

This is the modally strict principle which underlies the Fregean principle proved in (399)★.

(**395**) **Definition:** Directions. A *direction* is any object that is a direction of some line:

$$Direction(x) \equiv_{df} \exists u DirectionOf(x,u)$$

Cf. Frege 1884, §66.

(**396**) ★**Theorem:** Existence of Directions. If we assert, as modally fragile axiom, that lines exist ($\exists x Lx$), then it follows as a ★-theorem that directions exist:

$$\exists x Direction(x)$$

In Remark (390), we discussed taking $\exists y Ly$ as a modally fragile axiom. But it is not crucial – if this axiom is omitted, one can still prove, as a modally strict theorem, that $\exists x Lx \rightarrow \exists x Direction(x)$.

(**397**) ★**Theorem:** The Conditional Significance of The Direction of Line u.

$$\forall u(\imath x DirectionOf(x,u)\downarrow)$$

The present theorem also abbreviates a quantified conditional. Neither it nor the previous theorems require the existence of ordinary lines.

(**398**) **Definition:** Notation for The Direction of Line u. Though our conventions for using free restricted variables in definitions, developed in (339) weren't designed for *weak* restricted variables, they can be straightforwardly adapted without giving rise to logical problems. So we may introduce the following notation for the direction of line u:

$$\vec{u} =_{df} \imath x DirectionOf(x,u)$$

The expression $\vec{\kappa}$ is significant whenever κ is significant and known to be a line, either by hypothesis or by proof. But note that since the axiom or assumption that there exist lines ($\exists x Lx$) is modally fragile, one can't derive that the definiens in the above is significant by modally strict means. So the identities derived from this definition will be theorems, but not modally strict ones.

(**399**) ⋆**Theorem:** Fregean Biconditional for Directions. It now follows that for any two lines u and v, the direction of u is identical to the direction of v iff u and v are parallel:

$$\forall u \forall v(\vec{u}=\vec{v} \equiv u \| v)$$

We have therefore established the Fregean biconditional principle for directions (Frege 1884, §65).

(**400**) **Remark:** Pretheoretic Conception of Shapes. Consider the pretheoretic property *being a figure* (P) and the pretheoretic relation *being similar to*, where the latter is written $\sim$ in infix notation. Let us suppose, as in the case of lines, that it is a modally fragile axiom that there figures exist ($\exists xPx$), that Px is a weak restriction condition, and that $\sim$ is an equivalence relation among figures, i.e., where u, v are now weak restricted variables that range over figures, that

$$\forall u(u \sim u)$$

$$\forall u \forall v(u \sim v \rightarrow v \sim u)$$

$$\forall u \forall u' \forall u''(u \sim u' \,\&\, u' \sim u'' \rightarrow u \sim u'')$$

We may then apply our theory by abstracting over similar figures to define shapes.

(**401**) **Lemma:** Fact About Similarity. By the same reasoning we used in (391), we know that for any two figures u and v, the properties *being a figure similar to u* and *being a figure similar to v* are materially equivalent if and only if u is similar to v:

$$\forall u \forall u'(\forall z([\lambda v\, v \sim u]z \equiv [\lambda v\, v \sim u']z) \equiv u \sim u')$$

We now derive a Frege-style analysis of shapes.

(**402**) **Definition:** Shapes of Figures. Given the principles in Remark (400), we may define: x is a *shape of* figure u just in case x is an extension of the property *being a figure similar to u*:

$$\mathit{ShapeOf}(x,u) \equiv_{df} \mathit{ExtensionOf}(x,[\lambda v\, v \sim u])$$

(**403**) **Theorems:** The Conditional Existence of *ShapeOf*. It now follows that (.1) every figure has a shape; and (.2) every figure has a unique shape:

(.1) $\forall u \exists x \mathit{ShapeOf}(x,u)$

(.2) $\forall u \exists! x \mathit{ShapeOf}(x,u)$

(**404**) **Theorem:** Principle Underlying the Fregean Abstraction Principle for Shapes. If x is a shape of figure u and y is a shape of figure v, then x is identical to y if and only if u is similar to v:

$$(\mathit{ShapeOf}(x,u)\ \&\ \mathit{ShapeOf}(y,v)) \rightarrow (x = y \equiv u \sim v)$$

This is the modally strict principle which underlies the Fregean principle proved in (409)★.

(**405**) **Definition:** Shapes. A *shape* is any object that is a shape of some figure:

$$\mathit{Shape}(x) \ \equiv_{df} \ \exists u \mathit{ShapeOf}(x,u)$$

(**406**) ⋆**Theorem:** Existence of Shapes. It now follows, as a ⋆-theorem, that shapes exist:

$$\exists x \mathit{Shape}(x)$$

This theorem is analogous to theorem (396)★. If one prefers not to temporarily take $\exists x Px$ ('there exist shapes') as a modally fragile axiom, then it is still a modally strict theorem that $\exists x Px \rightarrow \exists x \mathit{Shape}(x)$.

(**407**) ⋆**Theorems:** The Conditional Significance of The Shape of u.

$$\forall u (\imath x \mathit{ShapeOf}(x,u)\downarrow)$$

Again, this theorem, (403.1), and (403.2) are quantified conditionals, given the presence of the restricted variable u.

(**408**) **Definition:** Notation for The Shape of a Figure u. We are therefore justified in introducing the following notation for the shape of figure u:

$$\tilde{u} \ =_{df} \ \imath x \mathit{ShapeOf}(x,u)$$

Again, this notation requires us to adapt our conventions for using free restricted variables in definitions, developed in (339), to *weak* restricted variables. Also, note again that the identities derivable from this definition are theorems, but not modally strict ones.

(**409**) ⋆**Theorem:** Fregean Biconditional for Shapes. It now follows that for any two figures u and v, the shape of u is identical to the shape of v iff u and v are similar:

$$\forall u \forall v (\tilde{u} = \tilde{v} \ \equiv \ u \sim v)$$

10.8.2 General Abstraction via Equivalence Relations

In this remainder of this chapter, we revert to using w as an unrestricted individual variable. Hence, in this final subsection, x,y,z,w are all unrestricted individual variables.

(**410**) **Definition:** Equivalence Relations. Let F be a binary relation variable. Then we say that F is an *equivalence* relation on individuals if and only if F is reflexive, symmetric, and transitive:

$$\begin{aligned}&\mathit{Equivalence}(F) \ \equiv_{df}\\ &\quad \forall x Fxx \ \&\ \forall x\forall y(Fxy \to Fyx) \ \&\ \forall x\forall y\forall z(Fxy \,\&\, Fyz \to Fxz)\end{aligned}$$

(**411**) **Theorem:** Example of an Equivalence Relation. It is straightforward to establish that the relation of *exemplifying the same properties* is an equivalence relation:

$$\mathit{Equivalence}([\lambda xy\, \forall F(Fx \equiv Fy)])$$

(**412**) **Metatheorem:** *Equivalence*(F) is a Restriction Condition. By the definition of a restriction condition in (336) and the fact that *Equivalence*(F) has a single free variable, the metatheorem is established by the following two facts:

(.1) $\vdash_{\Box} \exists F(\mathit{Equivalence}(F))$

(.2) $\vdash_{\Box} \mathit{Equivalence}(\Pi) \to \Pi\!\downarrow$, where Π is any binary relation term

Since *Equivalence*(F) is a restriction condition on relations, we henceforth use $\widetilde{F}$ as a restricted variable ranging over equivalence relations.

(**413**) **Lemmas:** Fact About Equivalence Relations. (.1) For any equivalence relation $\widetilde{F}$, the properties *bearing* $\widetilde{F}$ *to* x and *bearing* $\widetilde{F}$ *to* y are materially equivalent if and only if x bears $\widetilde{F}$ to y:

(.1) $\forall \widetilde{F}(\forall w([\lambda z\, \widetilde{F}zx]w \equiv [\lambda z\, \widetilde{F}zy]w) \equiv \widetilde{F}xy)$

Note also that (.2) for any equivalence relation $\widetilde{F}$, the individuals that bear $\widetilde{F}$ to x are precisely the individuals to which x bears $\widetilde{F}$, i.e.,

(.2) $\forall \widetilde{F}\forall y([\lambda z\, \widetilde{F}zx]y \equiv [\lambda z\, \widetilde{F}xz]y)$

As a fact about properties, (.2) says that $[\lambda z\, \widetilde{F}zx]$ is materially equivalent to $[\lambda z\, \widetilde{F}xz]$. Hence, (.3) for any binary equivalence relation $\widetilde{F}$, the properties materially equivalent to $[\lambda z \widetilde{F}zx]$ are precisely the properties materially equivalent to $[\lambda z\, \widetilde{F}xz]$, i.e.,

(.3) $\forall \widetilde{F}\forall G(\forall y(Gy \equiv [\lambda z\, \widetilde{F}zx]y) \equiv \forall y(Gy \equiv [\lambda z\, \widetilde{F}xz]y))$

(**414**) **Definition:** Abstractions from Equivalence Relations. Where $\widetilde{F}$ is any equivalence relation, we say that w is an $\widetilde{F}$*-abstraction* of x just in case w is an extension of the property $[\lambda z\,\widetilde{F}zx]$:

$$\widetilde{F}\textit{-AbstractionOf}\,(w,x) \equiv_{df} \textit{ExtensionOf}(w,[\lambda z\,\widetilde{F}zx])$$

(**415**) **Theorem:** Principle Underlying Frege's Abstraction Principle. If w is an $\widetilde{F}$-abstraction of x and z is an $\widetilde{F}$-abstraction of y, then w is identical to z if and only if x and y exemplify $\widetilde{F}$:

$$(\widetilde{F}\textit{-AbstractionOf}\,(w,x)\;\&\;\widetilde{F}\textit{-AbstractionOf}\,(z,y)) \rightarrow (w=z \equiv \widetilde{F}xy)$$

(**416**) **Theorems:** The Conditional Existence of $\widetilde{F}$-Abstractions. It now follows that for every equivalence relation $\widetilde{F}$, (.1) there is an $\widetilde{F}$-abstraction of individual x; and (.2) there is a unique $\widetilde{F}$-abstraction of individual x:

(.1) $\forall\widetilde{F}\exists w\widetilde{F}\textit{-AbstractionOf}(w,x)$

(.2) $\forall\widetilde{F}\exists!w\widetilde{F}\textit{-AbstractionOf}(w,x)$

These are all really quantified conditionals, given that $\widetilde{F}$ is a restricted variable ranging over equivalence relations.

(**417**) ⋆**Theorems:** The Conditional Significance of the $\widetilde{F}$-Abstraction of x. As with (397)⋆ and (407)⋆, we now have that (.3) the $\widetilde{F}$-abstraction of individual x exists:

$$\forall\widetilde{F}(\imath w\widetilde{F}\textit{-AbstractionOf}(w,x)\downarrow)$$

(**418**) **Definition:** Notation for the $\widetilde{F}$-Abstraction of x. Whenever $\widetilde{F}$ is an equivalence relation, we notate the $\widetilde{F}$-abstraction of x as follows:

$$\widehat{x}_{\widetilde{F}} =_{df} \imath w\widetilde{F}\textit{-AbstractionOf}(w,x)$$

Cf. Frege 1884, §68.

(**419**) ⋆**Theorem:** The Fregean Biconditional for Definition by Abstraction. To establish that the $\widetilde{F}$-abstraction of x is indeed defined by classical abstraction over an equivalence relation, we prove that it obeys the classical principle, namely, that for any equivalence relation $\widetilde{F}$, the $\widetilde{F}$-abstraction of x is identical to the $\widetilde{F}$-abstraction of y if and only if x bears $\widetilde{F}$ to y:

$$\forall\widetilde{F}(\widehat{x}_{\widetilde{F}}=\widehat{y}_{\widetilde{F}} \equiv \widetilde{F}xy)$$

Thus, the classical principle used for 'definitions by abstraction' falls out as a theorem. This theorem and the theorems in (389)⋆ *justify* the practice of definition by abstraction.[247]

[247] See Mancosu 2017 for a comprehensive history of this practice in mathematics.

Chapter 11

Platonic Forms

(**420**) **Remark:** Platonic Forms. Plato is well-known for having postulated *Forms* to explain why it is that, despite an ever-changing reality, we can truly say that *different* objects have something in common, such as when we say that distinct objects x and y are both red spheres, or beautiful paintings, virtuous persons, etc. Plato thought that since concrete objects are always undergoing change and have many of their characteristics only temporarily, there must be something that is universal and unchanging if we can truly say that different objects are both F. Plato called the aspects of reality that are universal and unchanging the *Forms* and supposed that objects acquire their characteristics by participating in, or partaking of, these Forms.

Plato's fundamental principle about the Forms is the One Over Many Principle. This principle is most famously stated in *Parmenides* 132a, and though we shall discuss it in more detail in what follows, a study of the seminal papers in Plato scholarship suggests that the following statement of the principle is accurate:[248]

(OM) **One Over Many Principle**
If x and y are both F, then there exists something that is the Form of F, or F-ness, in which they both participate.

In this principle, we are to substitute predicate nouns or adjectives for the symbol 'F', so that we have instances like the following:

> If x and y are both human, then there exists something that is the Form of humanity, or humanness, in which they both participate.

> If x and y are both red, then there exists something that is the Form of redness in which they both participate.

[248]See, e.g., Vlastos 1954, principles (A1) and (B1); Vlastos 1969, principle (1); and Strang 1963, principle (OM).

Though 'The Form of F' and 'F-ness' are two traditional ways of referring to the same thing, we must remember that in the expression 'F-ness', the letter 'F' is representing an arbitrary predicate noun or adjective and must be replaced by such an expression to produce a term denoting a Form. By contrast, in the theoretical/technical expression 'The Form of F', the symbol 'F' is a variable ranging over properties and so requires that 'F' be replaced by a gerund or abstract noun to yield a term denoting a Form.[249]

Now questions about (OM) immediately arise:

- What is *The Form of F*?
- What is participation?
- Is (OM) to be regarded as an axiom or can it be derived from more general principles?

To answer these questions, some philosophers have been tempted to identify *The Form of F* with the property F and to analyze: an object x *participates in*, or *partakes of*, The Form of F just in case x exemplifies F. We can formulate these analyses as follows:

(A) $\textit{The Form of } F \ =_{df}\ F$

(B) $\textit{ParticipatesIn}(x, F) \ =_{df}\ Fx$

Given (A) and (B), there is a natural formal representation of (OM):

(C) $(Fx \,\&\, Fy \,\&\, x \neq y) \rightarrow \exists G(G = F \,\&\, Gx \,\&\, Gy)$

This formula is a simple theorem of the second-order predicate calculus with identity.[250] So, (A), (B) and (C) provide answers, respectively, to the bulleted questions above.

Of course, some philosophers (e.g., the followers of Quine, nominalists, logical positivists, etc.) would object that the above analysis assumes second-order logic and its ontology of properties. But object theory has a theory of properties as rigorous as any mathematical theory and so provides a precise framework in which the above analysis can be put forward.

[249] For example, we may substitute 'human' for 'F' to produce the term 'humanness' (i.e., 'humanity') and substitute 'red' for 'F' to produce the term 'redness'. But we must substitute 'being human' or 'humanity' for 'F' to produce the term 'The Form of Being Human' or 'The Form of Humanity', and must substitute 'being red' or 'redness' to produce the term 'The Form of Being Red' or 'The Form of Redness'. A good rule of thumb is that if the symbol 'F' is being used in a natural language context, such as when we say 'x is F', then the symbol 'F' may be replaced by a predicate noun or adjective, but when the symbol 'F' is being used in a theoretical/technical or formal context, then the symbol 'F' is being used as a variable ranging over properties.

[250] Suppose $Fx \,\&\, Fy \,\&\, x \neq y$. Then by &E we have $Fx \,\&\, Fy$. Moreover, the laws of identity yield $F = F$. Hence, by &I, we obtain $F = F \,\&\, Fx \,\&\, Fy$. Hence, by $\exists$I, it follows that $\exists G(G = F \,\&\, Gx \,\&\, Gy)$.

Despite its precision, however, analysis (A) – (C) is problematic as a theory of Platonic Forms. We can begin to see why by considering a passage in Vlastos (1954), where he reformulates Plato's One Over Many Principle, as it occurs in *Parmenides* 132a1–b2 (1954, 320):

> This is the first step of the Argument, and may be generalized as follows:
>
> (A1) If a number of things, a, b, c are all F, there must be a single Form, F-ness, in virtue of which we apprehend a, b, c, as all F.
>
> Here 'F' stands for any discernible character or property. The use of the same symbol, 'F,' in 'F-ness,' the symbolic representation of the "single Form,"[5] records the identity of the character discerned in the particular ("large") and conceived in the Form ("Largeness") through which we see that this, or any other, particular has this character.

The footnote numbered 5 in this passage is rather interesting, for it includes the claim "That F and F-ness are logically and ontologically distinct is crucial to the argument" (1954, 320). So Vlastos is distinguishing the Form of F from the property F, though in other passages, he calls the property F "the predicative function of the same Form" (1954, note 39).

Geach (1956) also balks at the suggestion that, for Plato, the Form of F is the attribute or property F, at least in connection with Forms corresponding to 'kind terms' such as 'man' and 'bed'. He notes (1956, 74):

> Surely his [Plato's] chosen way of speaking of these Forms suggests that for him a Form was nothing like what people have since called an "attribute" or a "characteristic." The bed in my bedroom is to the Bed, not as a thing to an attribute or characteristic, but rather as a pound weight or yard measure in a shop to the standard pound or yard.

Geach thus takes the Form of F to be a "paradigm" individual that exemplifies F (1956, 76).[251] In his paper, he uses variables $x, y, \ldots$ to range over individuals

[251] In 1956, 76, Geach says that a *paradigm F* is something that *is F*; it is a *standard F* insofar as it is an exemplar of *F par excellence*. But Geach's suggestion is problematic; an alleged paradigmatic example of F has any number of properties that *undermine* that the suggestion that it is a paradigmatic exemplar. For suppose b is alleged to be a paradigm sphere. Then b will have a radius of a particular length l, be constructed of a particular material m, reflect light of a particular color c, etc. None of these properties are representative of spheres generally and the fact that b exemplifies these specific properties tends to undermine the claim that it is paradigmatic.

Intuitively, what Geach intends, is that the paradigm sphere, in so far as it is to serve as the Form of Sphericity, should be something abstracted from properties not strictly implied by being a sphere (like having a radius of a particular length l, being constructed of a particular material m, reflecting light of a particular color c), while exemplifying all and only the properties necessarily implied by *being a sphere*. This understanding of a *paradigm*, as something that 'has', in the sense of *encodes*, all and only the properties necessarily implied by *being a sphere*, will be captured by the 'thick' conception of the Forms discussed in the Section 11.2. Consequently, we postpone further discussion of this issue until then, and specifically until Remark (450).

(including Forms), and uses a different style of variable, namely $F, G, \ldots$, for properties (or attributes). Similarly, Strang (1963, 148) formulates a version of (OM) by using the notation 'A' to refer to the property A and '$F(A)$' to refer to a Form of A ("Given any set of A's, they participate in one and the same $F(A)$").

Allen (1960) offers a second reason why analysis (A) – (C) should not be adopted, namely, it doesn't make sense of a principle to which Plato often appeals:

(SP) **Self-Predication Principle**
The Form of F is F.

Allen explicitly notes (1960, 148):[252]

> Plato obviously accepts the following thesis: some (perhaps all) entities which may be designated by a phrase of the form "the *F* Itself," or any synonyms thereof, may be called *F*. So the Beautiful Itself will be beautiful, the Just Itself just, Equality equal.[3]

Allen assumed these facts in the following lines from the opening passage of his article (1960, 147):

> The significance—or lack of significance—of Plato's self-predicative statements has recently become a crux of scholarship. Briefly, the problem is this: the dialogues often use language which suggests that the Form is a universal which has itself as an attribute and is thus a member of its own class, and, by implication, that it is the one perfect member of that class. The language suggests that the Form *has* what it *is*: it is self-referential, self-predicable.
>
> Now such a view is, to say the least, peculiar. Proper universals are not instantiations of themselves, perfect or otherwise. Oddness is not odd; Justice is not just; Equality is equal to nothing at all. No one can curl up for a nap in the Divine Bedsteadity; not even God can scratch Doghood behind the Ears.
>
> The view is more than peculiar; it is absurd. . . .

With this, we are in a position to appreciate a later passage about (SP) in this same paper. Allen writes (1960, 148–9):

> But this thesis [SP] does not, by itself, imply self-predication; for that, an auxiliary premise is required. This premise is that a function of the type ". . . is *F*" may be applied univocally to *F* particulars and to the *F* Itself, so that when (for example) we say that a given act is just, and that Justice is just, we are asserting that both have identically the same character. But

[252] The footnote numbered 3 in the following quotation provides his documentation. Allen cites the following passages where Plato offers a version of the claim in question: *Protagoras*, 330c, 331b; *Phaedo*, 74b, d, 100c; *Hippias Major*, 289c, 291e, 292e, 294a-b; *Lysis*, 217a; *Symposium*, 210e-211d.

> this premise would be false if the function were systematically equivocal, according as the subject of the sentence was a Form or a particular. In that case, to say that Justice is just and that any given act is just would be to say two quite different (though perhaps related) things, and the difficulties inherent in self-predication could not possibly arise. ...I propose to show that functions involving the names of Forms exhibit just this kind of ambiguity.

Though I do not endorse many of the subsequent conclusions Allen draws in his paper, the conclusion in the above passage strikes me as insightful. The idea that there is ambiguity in predication and, indeed, that Plato saw the ambiguity, has been picked up by other Plato scholars, notably Frede (1967) and Meinwald (1992).[253]

Thus, the problem with the analysis (A) – (C) above is that it doesn't really make sense of Plato's text: it doesn't distinguish the property F from The Form of F and it can't make sense of the Self-Predication Principle. Vlastos' suggestion, that the property F and the Form of F are ontologically as well as logically distinct, and Allen's suggestion, that the language in the Self-Predication Principle involves a systematic ambiguity, are central to the analyses developed in this chapter.

Object theory distinguishes the property F from the abstract, logical individuals encoding F that might serve as *The Form of* F. Two such individuals immediately come to mind: (1) the abstract object that encodes just F and no other properties, and (2) the abstract object that encodes all and only the properties necessarily implied by F. The former offers a 'thin' conception on which The Form of F is the 'pure', objectified form of the property F ('thin' in the sense that it encodes a single property and 'pure' in the sense that it encodes no other property). The latter offers a 'thick' conception on which The Form of F encodes exactly what all F-exemplifiers necessarily exemplify in virtue of exemplifying F, i.e., those properties G such that $\Box\forall x(Fx \rightarrow Gx)$. The thin conception of Forms was developed in Zalta 1983 (Chapter II, Section 1), whereas the thick conception was developed in Pelletier & Zalta 2000. These conceptions are discussed below in some detail, in Sections 11.1 and 11.2, respectively. Though the thick conception may be a more considered and scholarly approach to Plato, the thin conception is not without interest, for it already yields theorems that provide plausible readings of (OM) and (SP).

[253] For a more complete history of Plato scholarship in which it is proposed that there is an ambiguity in predication, see the Appendix to Pelletier & Zalta 2000.

11.1 The Thin Conception of Forms

(**421**) **Definition:** A Thin Form of G. Intuitively, an individual x is a thin Form of property G if and only if x is an abstract object, G exists, and x encodes just the property G:

$$\mathit{ThinFormOf}(x,G) \equiv_{df} A!x \,\&\, G{\downarrow} \,\&\, \forall F(xF \equiv F = G)$$

Recall here that we are assuming Convention (17.2). Thus, where Π is an property term, $\mathit{ThinFormOf}(x,\Pi)$ is false, for any x, when $\neg\Pi{\downarrow}$.

(**422**) **Theorems:** There Exists a (Unique) Thin Form of G.

(.1) $\exists x(\mathit{ThinFormOf}(x,G))$

(.2) $\exists! x(\mathit{ThinFormOf}(x,G))$

(.3) $\imath x\mathit{ThinFormOf}(x,G){\downarrow}$

It also follows that if x is a thin Form of G and also a thin Form of H, then G is identical to H:

(.4) $\mathit{ThinFormOf}(x,G) \,\&\, \mathit{ThinFormOf}(x,H) \rightarrow G = H$

(**423**) **Definition:** Notation for the Thin Form of G. We may therefore introduce notation, $\boldsymbol{a}_G$, to designate the thin Form of G:

$$\boldsymbol{a}_G =_{df} \imath x\mathit{ThinFormOf}(x,G)$$

Exercise. Explain why the inferential role of the above definition-by-= guarantees, for any property term Π, that $\boldsymbol{a}_\Pi = \imath x\mathit{ThinFormOf}(x,\Pi)$ is a theorem if $\Pi{\downarrow}$, and $\neg\boldsymbol{a}_\Pi{\downarrow}$ is a theorem otherwise.

(**424**) **Theorem:** The Thin Form of G Is Canonical. It now follows that the thin Form of G is identical to the abstract object that encodes exactly the properties identical to G:

$$\boldsymbol{a}_G = \imath x(A!x \,\&\, \forall F(xF \equiv F = G))$$

So $\boldsymbol{a}_G$ is a canonical individual.

(**425**) **Remark:** $\boldsymbol{a}_G$ is Strictly Canonical. Given metadefinition (260.1), it can be shown that $F = G$ is a rigid condition on properties:

$$\vdash_\Box \forall F(F = G \rightarrow \Box F = G)$$

Given that $\boldsymbol{a}_G$ is canonical (424), the rigidity of the condition $F = G$ tells us that $\boldsymbol{a}_G$ is (identical to) a strictly canonical individual, as this was defined in (260.2). So the theorems in (261) apply to $\boldsymbol{a}_G$.

(**426**) **Theorems:** Facts About the Thin Form of G. It now follows, by modally strict proofs, that: (.1) the thin Form of G is abstract and encodes all and only the properties identical to G; and (.2) the thin Form of G is a thin Form of G:

(.1) $A!\boldsymbol{a}_G \,\&\, \forall F(\boldsymbol{a}_G F \equiv F = G)$

(.2) $\mathit{ThinFormOf}(\boldsymbol{a}_G, G)$

The interesting fact about (.2) is that it is derived without appeal to (145.2)★.

(**427**) **Theorem:** A Thin Form of G Encodes G. It is a simple consequence of definition (421) that if x is a thin Form of G, then x encodes G:

$$\mathit{ThinFormOf}(x, G) \rightarrow xG$$

(**428**) **Definition:** Participation. We now say that y *participates in* x if and only if there is a property F such that x is a thin Form of F and y exemplifies F:

$$\mathit{ParticipatesIn}(y, x) \;\equiv_{df}\; \exists F(\mathit{ThinFormOf}(x, F) \,\&\, Fy)$$

Cf. Zalta 1983 (42).[254] This definition will be refined in Section 11.2, where we discuss the thick conception of Forms and distinguish two kinds of participation corresponding to the two kinds of predication. But the above definition serves well enough for our purposes in the present section.

(**429**) **Lemma:** Thin Forms, Predication, and Participation. It is an immediate consequence of the previous definition that if x is a thin Form of G, then an individual y exemplifies G iff y participates in x:

$$\mathit{ThinFormOf}(x, G) \rightarrow \forall y(Gy \equiv \mathit{ParticipatesIn}(y, x))$$

(**430**) **Theorem:** The Equivalence of Exemplification and Participation. It is now a consequence that an object x exemplifies a property G iff x participates in the thin Form of G:

$$Gx \equiv \mathit{ParticipatesIn}(x, \boldsymbol{a}_G)$$

This theorem verifies that, under our analysis, Plato's notion of participation is equivalent to the modern notion of exemplification.

(**431**) **Theorem:** The One Over the Many Principle. As noted in Remark (420), Plato's most important principle governing the Forms is (OM): if there are two

[254] In Zalta 1983, $\mathit{ParticipatesIn}(y, x)$ was defined as: $\exists F(xF \,\&\, Fy)$. By reformulating the definition as we've done in here, an object participates only in those abstract objects that are thin Forms.

distinct individuals exemplifying G, then there exists something that is *the* Form of G in which they both participate. (OM) is validated by the following theorem:

$$Gx \,\&\, Gy \,\&\, x \neq y \rightarrow \exists z(z = \boldsymbol{a}_G \,\&\, ParticipatesIn(x,z) \,\&\, ParticipatesIn(y,z))$$

So we've preserved the main principle of Plato's theory without collapsing the distinction between the property G and the thin Form of G.

(**432**) **Theorems:** Facts About Thin Forms. Consider the property *being ordinary*, $O!$. It follows that (.1) the thin Form of G fails to exemplify $O!$:

(.1) $\neg O!\boldsymbol{a}_G$

Since (.1) holds for every G, by GEN, it follows that (.2) the thin Form of $O!$ fails to exemplify $O!$:

(.2) $\neg O!\boldsymbol{a}_{O!}$

(.2) will play a role in the Remark that follows the next theorem.

(**433**) **Theorem:** The Thin Form of G Encodes G and a Unique Property.

(.1) $\boldsymbol{a}_G G$

This is an encoding formula in which the individual term, $\boldsymbol{a}_G$, is complex. It also follows that the Thin Form of G encodes exactly one property:

(.2) $\exists! H \boldsymbol{a}_G H$

(**434**) **Remark:** The 'Self-Predication' Principle. Recall that (SP) was formulated in Remark (420) as "The Form of G is G". As we saw in that remark, Allen (1960) suggests that in Plato's work, the context "... is G" does not apply univocally to both individuals and the Form of G. Furthermore, as noted in Section 1.3, Meinwald (1992, 378) argues that the second half of Plato's *Parmenides* leads us "to recognize a distinction between two kinds of predication, marked ... by the phrases 'in relation to itself' (*pros heauto*) and 'in relation to the others' (*pros ta alla*)." (Intuitively, the Form of F *is* F in relation to itself, but ordinary things that exemplify F *are* F in relation to some other thing, namely, the Form of F.) Thus, Meinwald also takes (SP) to exhibit an ambiguity, and she suggests that it is true only if interpreted as a *pros heauto* predication.

Once we represent *pros heauto* predications as encoding predications and represent *pros ta alla* predications as exemplification predications, the ambiguity in (SP) can be resolved within the present system. The *pros heauto* reading is $\boldsymbol{a}_G G$, i.e., the thin Form of G encodes G. This is derivable as theorem (433.1) and hence true. The *pros ta alla* reading is $G\boldsymbol{a}_G$, i.e., the thin Form of G exemplifies G. We can prove in object theory that this is not generally true, since a counterexample is derivable. Indeed, we've already seen the counterexample, namely theorem (432), which asserts that $\neg O!\boldsymbol{a}_{O!}$. The thin Form of *being ordinary* fails to exemplify *being ordinary*. Hence $\neg\forall F(F\boldsymbol{a}_F)$.

Note also that $\neg G\boldsymbol{a}_G$ follows from the assumption that G is a concreteness-entailing property. A property F is *concreteness-entailing* just in case $\Box\forall x(Fx \to E!x)$. Intuitively, one might suppose that the following are such properties: *being extended in spacetime*, *being colored*, *having mass*, *being human*, etc. (Indeed, the property *being concrete* ($E!$) is provably concreteness-entailing.) So one can prove that if G is concreteness-entailing, then $\neg G\boldsymbol{a}_G$.[255]

Although we've now seen that the general form of (SP) is provably false when we represent the copula 'is' as *exemplification*, we shall see that there are special cases where, for some properties F, $\boldsymbol{a}_F$ does provably exemplify F.

(**435**) **Definition:** Thin Forms. We define: x is a *thin Form* if and only if x is a thin Form of F, for some F:

$$\mathit{ThinForm}(x) \equiv_{df} \exists F(\mathit{ThinFormOf}(x,F))$$

(**436**) **Theorem:** A Fact About Thin Forms. Clearly, the thin Form of G is a thin Form:

$$\mathit{ThinForm}(\boldsymbol{a}_G)$$

By GEN, this holds for any property G.

(**437**) **Theorems:** Facts About Thin Forms and Platonic Being. Suppose that instead of reading the defined term $A!$ as *being abstract*, we temporarily read it as *Platonic Being*. Then we can prove: (.1) Thin Forms exemplify Platonic Being (in ambiguous natural language, "Thin Forms are Platonic Beings"); (.2) thin Forms exemplify any property necessarily implied by Platonic Being; (.3) the thin Form of Platonic Being exemplifies Platonic Being; (.4) thin Forms participate in the thin Form of Platonic Being; and (.5) if F is necessarily implied by Platonic Being, then thin Forms participate in the thin Form of F:

(.1) $\mathit{ThinForm}(x) \to A!x$

(.2) $\Box\forall y(A!y \to Fy) \to (\mathit{ThinForm}(x) \to Fx)$

(.3) $A!\boldsymbol{a}_{A!}$

[255] Assume that P is a concreteness-entailing property and, for reductio, that $P\boldsymbol{a}_P$. Now since P is concreteness-entailing, it follows from the definition by the T-schema that $\forall x(Px \to E!x)$. Hence $P\boldsymbol{a}_P \to E!\boldsymbol{a}_P$. Since $P\boldsymbol{a}_P$ is our reductio assumption, it follows that $E!\boldsymbol{a}_P$. But by by (426.1), we know $A!\boldsymbol{a}_P$, which by definition (22.2) is the claim $[\lambda x\, \neg\Diamond E!x]\boldsymbol{a}_P$. By Rule $\overrightarrow{\beta}$C (184.1.a), it follows that $\neg\Diamond E!\boldsymbol{a}_P$, i.e., $\Box\neg E!\boldsymbol{a}_P$. By the T schema, this implies $\neg E!\boldsymbol{a}_P$, at which point we've reached a contradiction.

(.4) $\mathit{ThinForm}(x) \rightarrow \mathit{ParticipatesIn}(x, \boldsymbol{a}_{A!})$

(.5) $\Box\forall y(A!y \rightarrow Fy) \rightarrow (\mathit{ThinForm}(x) \rightarrow \mathit{ParticipatesIn}(x, \boldsymbol{a}_F))$

(**438**) **Theorems:** Thin Forms and Participation. (.1) There exists a thin Form that participates in itself; and (.2) there exists a thin Form that doesn't participate in itself:

(.1) $\exists x(\mathit{ThinForm}(x) \,\&\, \mathit{ParticipatesIn}(x, x))$

(.2) $\exists x(\mathit{ThinForm}(x) \,\&\, \neg \mathit{ParticipatesIn}(x, x))$

In case (.2) sounds familiar, we'll note here that in (440), we'll prove that *being a thin Form that doesn't participate in itself* doesn't exist; otherwise, it would be a Russell-style paradoxical property that implies a contradiction.

(**439**) **Remark:** The Third Man Argument. Plato puts forward an argument in *Parmenides* (132a) that has come to be known as the Third Man Argument (TMA). This argument raises a concern as to whether the theory of Forms involves an infinite regress. As Vlastos (1954, 321) develops the argument, Plato appears to draw an inference from (A1) to (A2), both of which are supposed to be instances of (OM):

(A1) If a number of things, a, b, c are all F, there must be a single Form, F-ness, in virtue of which we apprehend a, b, c, as all F.

(A2) If a, b, c, and F-ness are all F, there must be another Form, F_1-ness, in virtue of which we apprehend a, b, c, and F-ness as all F.

If this inference is valid and F_1-ness is distinct from F-ness, then it raises the concern that the inference is only the first step of a regress that commits Plato to an infinite number of Forms corresponding to the single property F. But Vlastos then notes (1954, 324–325) that for the inference from (A1) to (A2) to be valid, there seems to be two implicit assumptions, the Self-Predication Principle (SP) discussed earlier and the following:[256]

(NI) **Non-Identity Principle**
If x is F, x is not identical with The Form of F.

[256] Vlastos credits A. E. Taylor (1915-1916) for the suggestion that Plato needs the Self-Predication Principle. Indeed, if we examine Taylor 1915-1916, we find (253) that although that he accepts that it makes sense to predicate a universal of itself, he says:

> ... we must deny the tacit premiss of Parmenides that a universal can be predicated of itself as it is predicated of its instances.

So even as far back as 1915, Plato scholars recognized that by introducing two modes of predication, one can forestall the TMA.

Vlastos then observes (1954, 326) that these two tacit assumptions are jointly inconsistent. This was a trenchant observation, for if the two claims are formally represented as:

(SP′) $F\boldsymbol{a}_F$

(NI′) $Fx \rightarrow x \neq \boldsymbol{a}_F$

then the inconsistency becomes manifest; we can instantiate $\boldsymbol{a}_F$ into a universal generalization of (NI′) to obtain $F\boldsymbol{a}_F \rightarrow \boldsymbol{a}_F \neq \boldsymbol{a}_F$, and this, together with (SP′), yields the odious conclusion that $\boldsymbol{a}_F \neq \boldsymbol{a}_F$. Of course, the inconsistency of (SP′) and (NI′) involves a system in which there are: (a) principles that assert the existence and uniqueness of $\boldsymbol{a}_F$ and (b) principles governing (terms defined by) definite descriptions in general, and $\boldsymbol{a}_F$ in particular.

The literature that developed in response to Vlastos 1954 focused on (a) how to reformulate TMA so as to avoid the above inconsistency, and (b) whether there is textual support for the revised version of TMA. This literature included the papers Sellars 1955, Geach 1956, with rejoinders in Vlastos 1955 and 1956. Vlastos 1969 (footnote 2) includes a list of papers that were subsequently published on TMA, though see Cohen 1971, Meinwald 1992, and Pelletier & Zalta 2000 for further discussion and additional bibliography.

We conclude our discussion of TMA, as well as this section on the thin conception of the Forms, by noting that both of the tacit assumptions that Vlastos formulated for TMA, as he understood them, are provably false on the present theory. We've already seen that the above reading (SP′) of (SP) has a counterexample. Object theory also implies that the reading (NI′) of (NI) has a counterexample. By (437.3), we know $A!\boldsymbol{a}_{A!}$, and by the existence of $\boldsymbol{a}_{A!}$ and the reflexivity of identity (117.1), we know $\boldsymbol{a}_{A!} = \boldsymbol{a}_{A!}$. But the conjunction of these two conclusions, $A!\boldsymbol{a}_{A!}$ & $\boldsymbol{a}_{A!} = \boldsymbol{a}_{A!}$ is a counterexample to (NI′), as formally represented above. Hence, we shouldn't accept either of the two tacit assumptions that Vlastos thought were needed for TMA, at least not if the predication involved in (SP) and in the antecedent of (NI) are understood as exemplification predications. We'll return to the discussion of TMA at the end of Section 11.2, where we discuss the thick conception of the Forms.

(**440**) **Theorem:** An Empty λ-Expression. We now establish that (.1) *being a thin Form that doesn't participate in itself* doesn't exist, and (.2) there is no property F that is exemplified by all and only those objects x that are thin Forms that don't participate in themselves:

(.1) $\neg[\lambda x\, \mathit{ThinForm}(x) \,\&\, \neg \mathit{ParticipatesIn}(x,x)]\downarrow$

(.2) $\neg\exists F \forall x(Fx \equiv \mathit{ThinForm}(x) \,\&\, \neg \mathit{ParticipatesIn}(x,x))$

So the property one might try to formulate to derive a Russell-style paradox from Plato's Theory of Forms doesn't exist.

11.2 The Thick Conception of Forms

(**441**) **Remark:** The Need for a Thick Conception of the Forms.[257] We've seen how our work thus far validates (a) the view that truly predicating F of a Form of F is different from truly predicating F of ordinary things, and (b) Meinwald's (1992) idea that in the *Parmenides*, Plato is attempting to get his audience (i) to appreciate a distinction in predication, namely, between saying that x is G *pros heauto* (i.e., in relation to itself) and saying that x is G *pros ta alla* (i.e., in relation to the others),[258] and (ii) to recognize that the Self-Predication principle (SP) is generally true only if read as a *pros heauto* predication. We've also seen that, in object theory, where encoding predication formally represents *pros heauto* predication and the Form of G is conceived as the 'thin' abstract object that encodes just the single property G, (SP) is preserved as theorem (433.1), which asserts that the Form of G encodes G.

In what follows, however, we consider some new data that our analysis thus far doesn't accommodate. According to Meinwald, there are also true *pros heauto* predications such as "The Form of G is F", where F is a property distinct from G. Meinwald lists the following such claims as ones Plato would endorse (1992, 379):[259]

> The Just is virtuous (*pros heauto*).
> Triangularity is three-sided (*pros heauto*).
> Dancing moves (*pros heauto*).

She explains these as follows (1992, 379–80):[260]

> The Just is Virtuous

[257] I'm indebted to F. Jeffry Pelletier for his critical contributions to Pelletier & Zalta 2000. The thick conception of Forms was first developed in that paper; Jeff recognized that the Plato scholars had developed a view of the Forms that called for a thick conception. Many of the ideas in what follows were first developed in that paper.

[258] We remind the reader again that Meinwald (1992, 381) relates the two kinds of predication in the *Parmenides* to the *kath' hauto* and *pros allo* uses of 'is' in the *Sophist*, as these are distinguished in Frede 1967, 1992.

[259] Meinwald writes:

> It is clear that such sentences come out true in Plato's work, as well as fitting our characterizations of predication of a subject in relation to itself.

Unfortunately, though, Meinwald doesn't identify any passages where Plato discusses examples like the ones to follow.

[260] In the following quotation, Meinwald refers to a *tree predication*. The *tree* in question is a genus-species structure, in which the characteristics associated with a genus are inherited by, and thus predicable of, the members of the species covered by the genus. Thus, a true tree predication is one in which a member of a species is F in virtue of the fact that F is one of the characteristics associated with the covering genus. In Meinwald's example, since dancing is a species of motion, 'Dancing moves' is a true tree predication.

> holds because of the relationship between the natures associated with its subject and predicate terms: Being virtuous is part of what it is to be just. Or we can describe predication as holding because Justice is a kind of Virtue. If we assume that to be a triangle is to be a three-sided plane figure (i.e., that Triangle is the species of the genus Plane Figure that has the differentia Three-Sided), then
>
> > Triangularity is three-sided.
>
> holds too. We can also see that
>
> > Dancing moves.
>
> is a true tree predication, since Motion figures in the account of what Dancing is.

Clearly, if the Form of G is identified as it was in the preceding section, i.e., as the abstract object that encodes G and no other property, then we cannot as yet interpret the above *pros heauto* predications as encoding predications about the Forms in question. For on such a 'thin' conception, the Form of Justice, $\boldsymbol{a}_J$, encodes only the property *being just* (J) and so does not encode the distinct property *being virtuous* (V). Similarly, the Form of Triangularity ($\boldsymbol{a}_T$) encodes only the property *being triangular* (T) and not the distinct property *being three-sided* ($3S$), and so isn't three-sided *pros heauto*. And similarly for the Form of Dancing.

However, on the thick conception, the Form of G is the abstract object that encodes all the properties necessarily implied by G, i.e., encodes all and only the properties F such that, necessarily, anything that exemplifies G exemplifies F. In what follows, we shall introduce the functional term Φ_G to designate this thick Form of G. So if the Form of Justice (Φ_J) encodes all of the properties necessarily implied by *being just*, then from the premise that *being virtuous* (V) is necessarily implied by *being just*, it follows that Φ_J encodes V. Thus, the encoding claim $\Phi_J V$ ("the Form of Justice encodes being virtuous") provides a reading of the *pros heauto* predication "The Just is virtuous".

Similarly, from the facts that:

> *Being triangular* necessarily implies *being three-sided*.
>
> *Dancing* necessarily implies *being in motion*.

it follows, on the thick conception of the Forms of Triangularity (Φ_T) and Dancing (Φ_D), that:

> The Form of Triangularity encodes *being three-sided*.
> $\Phi_T 3S$

The Form of Dancing encodes *being in motion*.
$\Phi_D M$

So the thick conception of the Forms is distinguished from the thin conception by the fact that it can provide an analysis of *pros heauto* predications of the form "The Form of G is F" when supplemented by facts about property implication and property distinctness. Under the thin conception, we cannot derive $\boldsymbol{a}_G F$ given the premises that F is distinct from G and necessarily implied by G. Under the thick conception, we can derive $\Phi_G F$ from G necessarily implies F; this is a consequence of theorem (449.2) established below. *A fortiori*, we can derive $\Phi_G F$ from the conjunction of $F \neq G$ and G necessarily implies F. These derivations, we suggest, are the best way to understand and represent the conditions under which "The Form of G is F *pros heauto*" is true.

(**442**) **Definition:** Property Implication and Equivalence. To define the thick conception of Plato's Forms, we need some preliminary definitions. We first define G *necessarily implies* F, written '$G \Rightarrow F$', as: necessarily, everything that exemplifies G exemplifies F:

(.1) $G \Rightarrow F \ \equiv_{df} \ F\downarrow \;\&\; G\downarrow \;\&\; \Box\forall x(Gx \rightarrow Fx)$

Thus, the claim *being triangular* necessarily implies *being three-sided* is defined to mean that *being triangular* and *being three-sided* both exist, and necessarily, any object exemplifying the former exemplifies the latter. In the usual manner, we may not formulate instances of this definition by uniformly replacing G or F with property terms in which x occurs free.

Clearly, we need to add the existence clauses to the definiens so that a claim of the form $\Pi \Rightarrow \Pi'$ will be provably false if either Π or Π' fails to be significant. Without the existence clause $\Pi\downarrow$ in the definiens, the claim $\Pi \Rightarrow \Pi'$ would be provably true when $\neg\Pi\downarrow$, since it would immediately follow that $\neg\Pi x$ by axiom (39.5.a), thereby implying $\Pi x \rightarrow \Pi' x$, which by GEN and RN yields $\Box\forall x(\Pi x \rightarrow \Pi' x)$. Without the existence clause $\Pi'\downarrow$ in the definiens, the claim $\Pi \Rightarrow \Pi'$ would be provably true when $\neg\Pi'\downarrow$, $\Pi\downarrow$, and Π is necessarily unexemplified (i.e., $\Box\forall x\neg\Pi x$), for in that case, it would again follow that $\neg\Pi x$ (this time by the T schema and $\forall$E), again implying $\Pi x \rightarrow \Pi' x$ and yielding $\Box\forall x(\Pi x \rightarrow \Pi' x)$.

We also say that properties G and F are *necessarily equivalent* just in case they necessarily imply each other:

(.2) $G \Leftrightarrow F \ \equiv_{df} \ G \Rightarrow F \;\&\; F \Rightarrow G$

Of course, when reasoning with (.1) and (.2), we can ignore the existence clauses as long as significant terms are in play.

(**443**) **Theorems:** Facts About Necessary Equivalence. As simple consequences of the preceding definitions, we have (.1) G and F are necessarily equivalent just in case necessarily, all and only the individuals exemplifying G exemplify F; (.2) necessarily equivalence is a reflexive, symmetric, and transitive condition; and (.3) properties are necessarily equivalent if and only they necessarily imply the same properties:

(.1) $G \Leftrightarrow F \equiv \Box\forall x(Gx \equiv Fx)$

(.2) $\Leftrightarrow$ is an equivalence condition:

(.a) $G \Leftrightarrow G$ ($\Leftrightarrow$ is reflexive)

(.b) $G \Leftrightarrow F \rightarrow F \Leftrightarrow G$ ($\Leftrightarrow$ is symmetric)

(.c) $(G \Leftrightarrow F \,\&\, F \Leftrightarrow H) \rightarrow G \Leftrightarrow H$ ($\Leftrightarrow$ is transitive)

(.3) $G \Leftrightarrow F \equiv \forall H(G \Rightarrow H \equiv F \Rightarrow H)$

Moreover, recall that by definition (200.1), a property F is necessary just in case $\Box\forall xFx$, and by definition (200.2), a property F is impossible just in case $\Box\forall x\neg Fx$. Then we have: (.4) if G and F are necessary properties, then G and F are necessarily equivalent; and (.5) if G and F are impossible properties, then G and F are necessarily equivalent:

(.4) $(\mathit{Necessary}(G) \,\&\, \mathit{Necessary}(F)) \rightarrow G \Leftrightarrow F$

(.5) $(\mathit{Impossible}(G) \,\&\, \mathit{Impossible}(F)) \rightarrow G \Leftrightarrow F$

(**444**) **Definition:** A (Thick) Form of G. We define: x is a *Form of* G iff x is an abstract object, G exists, and x encodes all and only the properties necessarily implied by G:

$$\mathit{FormOf}(x,G) \equiv_{df} A!x \,\&\, G{\downarrow} \,\&\, \forall F(xF \equiv G \Rightarrow F)$$

Note that the existence clause $G{\downarrow}$ is still needed in the definiens, notwithstanding its presence in the definiens of $G \Rightarrow F$. Without it, one could prove that $\mathit{FormOf}(x,\Pi)$ when Π is an empty property term and x is assigned the null object.[261]

(**445**) **Theorems:** There Exists a (Unique) Form of G.

[261] Suppose we had defined: $\mathit{FormOf}(x,G) \equiv_{df} A!x \,\&\, \forall F(xF \equiv G \Rightarrow F)$. Then following theorem would be licensed by the definition, where $\boldsymbol{a}_\varnothing$ is the null object and Π is any empty property term:

$$\mathit{FormOf}(\boldsymbol{a}_\varnothing,\Pi) \equiv (A!\boldsymbol{a}_\varnothing \,\&\, \forall F(\boldsymbol{a}_\varnothing F \equiv \Pi \Rightarrow F))$$

Now by definition of $\Rightarrow$, it is a modally strict theorem that:

$$\Pi \Rightarrow F \equiv (\Pi{\downarrow} \,\&\, F{\downarrow} \,\&\, \Box\forall z(\Pi z \rightarrow Fz))$$

So it would follow by a rule of substitution that:

(ϑ) $\mathit{FormOf}(\boldsymbol{a}_\varnothing,\Pi) \equiv (A!\boldsymbol{a}_\varnothing \,\&\, \forall F(\boldsymbol{a}_\varnothing F \equiv \Pi{\downarrow} \,\&\, F{\downarrow} \,\&\, \Box\forall z(\Pi z \rightarrow Fz)))$

(.1) $\exists x FormOf(x,G)$

(.2) $\exists! x FormOf(x,G)$

(.3) $\imath x FormOf(x,G)\downarrow$

(**446**) **Definition:** Notation for the Form of G. Given the previous theorem, we introduce the notation Φ_G to designate The Form of G:

$$\Phi_G =_{df} \imath x FormOf(x,G)$$

(**447**) **Theorem:** The Form of G is Canonical. It follows that the Form of G is identical to the abstract object that encodes all and only the properties necessarily implied by G:

$$\Phi_G = \imath x(A!x \,\&\, \forall F(xF \equiv G \Rightarrow F))$$

(**448**) **Remark:** The Form of G is Strictly Canonical. Note that the condition $G \Rightarrow F$ is a rigid condition on properties, as this was defined in (260.1):

$$\vdash_{\Box} \forall F(G \Rightarrow F \rightarrow \Box G \Rightarrow F)$$

Since Φ_G is canonical (447) and $G \Rightarrow F$ is a rigid condition on properties, we know that Φ_G is (identical to) a strictly canonical individual, as this was defined in (260.2). So the theorems in (261) apply to Φ_G.

(**449**) **Theorems:** Facts About Φ_G. It now follows, by modally strict reasoning, that: (.1) The Form of G is abstract and encodes exactly the properties F necessarily implied G; and (.2) The Form of G is a Form of G:

(.1) $A!\Phi_G \,\&\, \forall F(\Phi_G F \equiv G \Rightarrow F)$

(.2) $FormOf(\Phi_G, G)$

It follows *a fortiori* from the second conjunct of (.1) that $(G \Rightarrow F) \rightarrow \Phi_G F$. So by (63.10), $G \Rightarrow F \vdash \Phi_G F$. And by (63.7), $F \neq G$, $G \Rightarrow F \vdash \Phi_G F$. Thus, our theorems validate the distinguishing feature of the thick conception of the Forms discussed at the end of Remark (441).

(**450**) **Remark:** Thick Forms as Paradigms. It is worth pausing to note the sense in which the thick conception of Forms validates Geach's suggestion that

---

One could then derive $FormOf(\boldsymbol{a}_{\varnothing},\Pi)$ by deriving the right-side of (ϑ). Clearly, $A!\boldsymbol{a}_{\varnothing}$. So, by GEN, it remains to show:

(ξ) $\boldsymbol{a}_{\varnothing}F \equiv \Pi\downarrow \,\&\, F\downarrow \,\&\, \Box\forall z(\Pi z \rightarrow Fz)$

But, this is easy to establish, since both sides are provably false. It is a fact that $\neg\boldsymbol{a}_{\varnothing}F$, by (266.3) and (263.1). And since $\neg\Pi\downarrow$ by assumption, the right-side of (ξ) is false.

So, without the clause $G\downarrow$ in the definiens of *FormOf*, we could prove $FormOf(\boldsymbol{a}_{\varnothing},\Pi)$ when $\neg\Pi\downarrow$.

Forms are *paradigms*. Let S denote the ordinary property *being spherical* that we use pre-theoretically to characterize concrete objects, and consider the Form of Being Spherical, Φ_S. On the thick conception, Φ_S encodes all and only those properties necessarily implied by *being spherical*. Thus, Φ_S encodes *having a radius of some particular length*, since that is implied by *being spherical*, although there isn't a particular length l such that Φ_S encodes *being of length l*. Similarly, Φ_S encodes *being constructed of some particular material*, since that is implied by *being spherical*, but there won't be a particular material m such that Φ_S encodes *being constructed of material m*. And so on. Properties not strictly implied by G aren't encoded by the (thick) Form of G. The above conception avoids the problem, discussed in footnote 251, that afflicts Geach's suggestion that the Form of G is a paradigm *exemplifier* of G, since a paradigm G-exemplifier exemplifies properties not implied by G. And, clearly, since encoding is a mode of predication, there is a sense in which the (thick) Form of G *has* just those properties necessarily implied by G.

(**451**) **Definitions:** Two Kinds of Participation. We may introduce two kinds of participation to correspond to the two kinds of predication. We say: (.1) an object y participates *pros ta alla* in x (written $ParticipatesIn_{\text{PTA}}(y,x)$) iff there is a property F such that x is a Form of F and y exemplifies F, and (.2) an object y participates *pros heauto* in x (written $ParticipatesIn_{\text{PH}}(y,x)$) iff there is a property F such that x is a Form of F and y encodes F:

(.1) $ParticipatesIn_{\text{PTA}}(y,x) \equiv_{df} \exists F(FormOf(x,F) \,\&\, Fy)$

(.2) $ParticipatesIn_{\text{PH}}(y,x) \equiv_{df} \exists F(FormOf(x,F) \,\&\, yF)$

(**452**) **Lemmas:** Forms, Predication, and Participation. It is an immediate consequence of the previous definitions that (.1) if x is a Form of G, then every individual y is such that y exemplifies G if and only if y participates$_{\text{PTA}}$ in x; (.2) if x is a Form of G, then every individual y is such that if y encodes G, then y participates$_{\text{PH}}$ in x:

(.1) $FormOf(x,G) \rightarrow \forall y(Gy \equiv ParticipatesIn_{\text{PTA}}(y,x))$

(.2) $FormOf(x,G) \rightarrow \forall y(yG \rightarrow ParticipatesIn_{\text{PH}}(y,x))$

(**453**) **Remark:** Why the Consequent of (452.2) is not a Quantified Biconditional. One might wonder why the consequent of (452.2) is just a quantified conditional and not a quantified biconditional. Specifically, when $FormOf(x,G)$, why doesn't $ParticipatesIn_{\text{PH}}(y,x)$ imply yG? The key fact behind an answer to this question is that the Forms of distinct, but necessarily equivalent, properties are identical to one another, and this creates the conditions for a counterexample to:

(ϑ) $FormOf(x,G) \rightarrow \forall y(ParticipatesIn_{\text{PH}}(y,x) \rightarrow yG)$

This was reported in Fitelson & Zalta 2007.[262] An extended discussion is required to fully document such a counterexample.

Consider the following two properties:

$$P = [\lambda x\, Qx \,\&\, \neg Qx] \qquad (Q \text{ any property})$$

$$T = [\lambda x\, Bx \,\&\, \forall y(Sxy \equiv \neg Syy)] \qquad (B \text{ any property, } S \text{ any relation})$$

If, say, Q is *being round*, B is *being a barber*, and S is the relation *x shaves y*, then P is the property *being round and not round* while T is the property *being a barber who shaves all and only those individuals that don't shave themselves*. It is reasonable to assert that these are distinct properties, i.e., that $P \neq T$.[263] But note that P and T are both provably *impossible* properties, for it is straightforward to show:

$$\Box\forall x\neg Px$$

$$\Box\forall x\neg Tx$$

Hence by (443.5), it follows that P and T are necessarily equivalent, and so by (443.3), that P and T necessarily imply the same properties, i.e.,

(ζ) $\forall F(P \Rightarrow F \;\equiv\; T \Rightarrow F)$

Now consider the following two instances of theorem (445.1):

$$\exists x FormOf(x,P)$$

$$\exists x FormOf(x,T)$$

Let b and c be such objects, respectively, so that we know $FormOf(b,P)$ and $FormOf(c,T)$. By definition (444), we know all of the following, among other things:

(A) $A!b$

(B) $\forall F(bF \;\equiv\; P \Rightarrow F)$

[262] The counterexample described in that paper was discovered computationally, using PROVER9. The counterexample corrected an error in Pelletier & Zalta 2000, which had mistakenly suggested that one could prove (452.2) when the connective in the consequent is strengthened to a biconditional.

[263] There are various *pretheoretic* reasons one might give for this: (a) the informal argument that shows an object couldn't possibly exemplify P must appeal to the property Q rather than to the property B and the relation S, whereas the informal argument that shows an object couldn't possibly exemplify T must appeal to B and S rather than to Q; and (b) one can tell a story about an object that is T without thereby telling a story about an object that is P.

(C) $A!c$

(D) $\forall F(cF \equiv T \Rightarrow F)$

Despite the fact that $P \neq T$, it can be shown that b and c are identical, and this is crucial to the construction of our counterexample. To show that b and c are identical, it suffices, by (A), (C), and (245), to show $\forall F(bF \equiv cF)$. But this last claim follows straightforwardly from (ζ), (B) and (D). Hence by definition of abstract object identity, we've established:

(G) $b = c$

Now to see how this leads to a counterexample to (ϑ), consider the fact that by the Comprehension Principle for Abstract Objects (53), there is an abstract object that encodes just T and no other properties:

$$\exists x(A!x \,\&\, \forall F(xF \equiv F = T))$$

This is, of course, the thin Form of T, but let's not get distracted by that at the moment. Let d be such an object, so that we know:

(H) $A!d \,\&\, \forall F(dF \equiv F = T)$

Now we can show that the following elements of the counterexample to (ϑ) are all true:

(i) $\mathit{FormOf}(b, P)$

(ii) $\mathit{ParticipatesIn}_{\text{PH}}(d, b)$

(iii) $\neg dP$

(i) is already known. To show (ii), i.e., $\mathit{ParticipatesIn}_{\text{PH}}(d, b)$, we have to show:

$$\exists F(\mathit{FormOf}(b, F) \,\&\, dF)$$

By $\exists$I, it suffices to show $\mathit{FormOf}(b, T) \& dT$. But $\mathit{FormOf}(c, T)$ is already known, and from (G) we know $b = c$. Hence $\mathit{FormOf}(b, T)$, by Rule =E. Moreover, dT follows immediately from (H), by instantiating the second conjunct of (H) to T and applying the reflexivity of identity. So $\exists F(\mathit{FormOf}(b, F) \,\&\, dF)$. It remains to show (iii), i.e., $\neg dP$. Note that from the right conjunct of (H), it follows that $dP \equiv P = T$. But by hypothesis, $P \neq T$. Hence, $\neg dP$.

Thus, we've established the elements of the counterexample to (ϑ). Consequently, under the reasonable hypothesis that there are distinct, impossible properties, $\mathit{ParticipatesIn}_{\text{PH}}(y, x)$ doesn't imply yG when $\mathit{FormOf}(x, G)$.

> **Exercise:** Show that $\mathit{ThinFormOf}(x, G) \to \forall y(yG \equiv \mathit{ParticipatesIn}(y, x))$ is a theorem when $\mathit{ThinFormOf}$ is defined as in (421) and $\mathit{ParticipatesIn}$ is defined as in (428).

(**454**) **Theorems:** Exemplification, Participation$_{\text{PTA}}$, Encoding, and Participation$_{\text{PH}}$. It is now a consequence that (.1) x exemplifies G iff x participates$_{\text{PTA}}$ in The Form of G; and (.2) if x encodes G, then x participates$_{\text{PH}}$ in the Form of G:

(.1) $Gx \equiv ParticipatesIn_{\text{PTA}}(x, \Phi_G)$

(.2) $xG \rightarrow ParticipatesIn_{\text{PH}}(x, \Phi_G)$

The discussion in Remark (453) explains why (.2) is a conditional and not a biconditional. Note also that these theorems are modally strict, despite the fact that they are conditionals in which a term defined by a rigid definite description appears in one of the conditions.

(**455**) **Theorem:** *ParticipatesIn*$_{\text{PTA}}$ Fact. It is a consequence of the foregoing definitions that if y participates$_{\text{PTA}}$in x, then y exemplifies every property x encodes:

$$ParticipatesIn_{\text{PTA}}(y, x) \rightarrow \forall F(xF \rightarrow Fy)$$

(**456**) **Theorems:** Two Versions of the One Over Many Principle. (OM) may now be derived in two forms: (.1) if there are two distinct individuals exemplifying G, then there exists something that is (identical to) *the* Form of G in which they both participate$_{\text{PTA}}$; (.2) if there are two distinct individuals encoding G, then there exists something that is (identical to) *the* Form of G in which they both participate$_{\text{PH}}$:

(.1) $Gx \,\&\, Gy \,\&\, x \neq y \rightarrow \exists z(z = \Phi_G \,\&\, ParticipatesIn_{\text{PTA}}(x, z) \,\&\, ParticipatesIn_{\text{PTA}}(y, z))$

(.2) $xG \,\&\, yG \,\&\, x \neq y \rightarrow \exists z(z = \Phi_G \,\&\, ParticipatesIn_{\text{PH}}(x, z) \,\&\, ParticipatesIn_{\text{PH}}(y, z))$

So we've preserved the main principle of Plato's theory not only in a version that governs exemplification and participation$_{\text{PTA}}$, but also in a version that governs encoding and participation$_{\text{PH}}$.

(**457**) **Theorems:** Counterexample to One Version of SP. Under the thick conception of the Forms, there are counterexamples to the principle that "The Form of F is F" when 'is' is interpreted as exemplification. Consider the property *being ordinary*, $O!$. It follows that (.1) the Form of G fails to exemplify $O!$; (.2) the Form of $O!$ fails to exemplify $O!$; and (.3) for some property G, the Form of G fails to exemplify G:

(.1) $\neg O!\Phi_G$

(.2) $\neg O!\Phi_{O!}$

(.3) $\exists G \neg G\Phi_G$

The first two are the counterparts of theorems (432.1) and (432.2) governing the thin conception of Forms. (.3) can be stated in an equivalent form: it isn't universally the case that Φ_G is G *pros ta alla*. Thus, an observation similar to the one discussed in Remark (434) holds for the thick conception of Forms: (457.3) establishes that the exemplification reading of (SP) fails to be universally true.

(**458**) **Theorem:** Version of SP that is Provable. When we interpret the 'is' in "The Form of F is F" as encodes, the result "The Form of F encodes F is a theorem:

$$\Phi_G G$$

As Plato would put it, The Form of G is G *pros heauto*. Thus, the reading of the Self-Predication Principle (SP) discussed in (433.1) also holds for Φ_G.

(**459**) **Definition:** Forms. Forms, under the thick conception, are defined in the same general way as under the thin conception: x is a *Form* if and only if x is a Form of G, for some G:

$$Form(x) \equiv_{df} \exists G(FormOf(x,G))$$

(**460**) **Theorems:** Some Facts About Forms. (.1) The Form of G is a Form; (.2) there exists a Form that doesn't participate$_{\text{PTA}}$ in itself; and (.3) every Form participates$_{\text{PH}}$ in itself:

(.1) $Form(\Phi_G)$

(.2) $\exists x(Form(x) \,\&\, \neg ParticipatesIn_{\text{PTA}}(x,x))$

(.3) $\forall x(Form(x) \rightarrow ParticipatesIn_{\text{PH}}(x,x))$

As an exercise, explain why the existence of Forms that don't participate$_{\text{PTA}}$ in themselves doesn't give rise to paradox.

(**461**) **Theorems:** Facts About 'Self-Predication' *Pros Ta Alla* and Self-Participation. Again, if we suppose that the property *being abstract* is the property *Platonic Being*, then we have the following theorems: (.1) The Form of *Platonic Being* exemplifies *Platonic Being*; (.2) The Form of *Platonic Being* participates$_{\text{PTA}}$ in itself; (.3) there exists a Form that participates$_{\text{PTA}}$ in itself; (.4) if *Platonic Being* necessarily implies H, then for every property G, The Form of Φ_G exemplifies H; hence, (.5) if *Platonic Being* necessarily implies H, then The Form of Φ_H exemplifies H. Moreover, (.6) if *Platonic Being* necessarily implies the negation of H, then for every property G, $\neg H\Phi_G$, and hence (.7) if *Platonic Being* necessarily implies the negation of H, then $\neg H\Phi_H$:

(.1) $A!\Phi_{A!}$

(.2) $ParticipatesIn_{\text{PTA}}(\Phi_{A!}, \Phi_{A!})$

(.3) $\exists x(Form(x) \mathbin{\&} ParticipatesIn_{\text{PTA}}(x, x))$

(.4) $(A! \Rightarrow H) \rightarrow \forall G(H\Phi_G)$

(.5) $(A! \Rightarrow H) \rightarrow H\Phi_H$

(.6) $(A! \Rightarrow \overline{H}) \rightarrow \forall G(\neg H\Phi_G)$

(.7) $(A! \Rightarrow \overline{H}) \rightarrow \neg H\Phi_H$

Theorem (.1) is a special case where a Form, namely The Form of Platonic Being, unconditionally exemplifies its defining property *pros ta alla*. This immediately implies (.2), that this Form participates$_{\text{PTA}}$ in itself. (.3) is then an immediate consequence of (.2). As an example of (.4), we now know that if *Platonic Being* necessarily implies *being at rest*, then for every property G, Φ_G exemplifies *being at rest*. In particular, (.5) if *Platonic Being* necessarily implies *being at rest* (R), then The Form of *being at rest*, Φ_R, exemplifies *being at rest*. This constitutes a conditional 'self-predication' *pros ta alla*. Note also that theorem (.7) is a simple consequence of (.6); as an example of (.7), if *Platonic Being* necessarily implies *not being extended* ($\overline{E}$), then The Form of *being extended*, Φ_E, fails to exemplify *being extended*.

(**462**) **Theorems:** Necessary Implication and Participation.

(.1) $(A! \Rightarrow H) \rightarrow \forall G(ParticipatesIn_{\text{PTA}}(\Phi_G, \Phi_H))$

(.2) $(G \Rightarrow H) \rightarrow ParticipatesIn_{\text{PH}}(\Phi_G, \Phi_H)$

As an example of (.1), if *being abstract* necessarily implies *being at rest* (R), then for any G, The Form of G participates$_{\text{PTA}}$ in The Form of Rest Φ_R. As an example of (.2), if *being just* (J) necessarily implies *being virtuous* (V), then The Form of Justice, Φ_J, participates$_{\text{PH}}$ in The Form of Virtue, Φ_V; and if *being a triangle* (T) necessarily implies *being three-sided* ($3S$), then Φ_T participates$_{\text{PH}}$ in Φ_{3S}. These last examples of (.2) show that given certain modal facts about properties, the data that Meinwald adduces for a thick conception of Forms, described in (441), become derivable.

(**463**) **Remark:** Platonic Analysis and Derivation of a Syllogism. Our definitions of participation and our definition of The Form of G offers a Platonic analysis of a classic form of syllogism. Consider:

Humans are mortal.
Socrates is a human.
———
Socrates is mortal.

On a Platonic analysis of this argument, the conclusion is validly derivable from the premises. Both the minor premise ('Socrates is a human') and the conclusion can be analyzed as asserting that Socrates participates$_{\text{PTA}}$ in a certain Form. The major premise ('Humans are mortal') can be analyzed as asserting that The Form of Humanity is mortal *pros heauto*. Formally, where 's' denotes Socrates, 'H' denotes *being human*, and 'M' denotes *being mortal*, these analyses can be captured as follows:

$$\begin{array}{l} \Phi_H M \\ \mathit{ParticipatesIn}_{\text{PTA}}(s, \Phi_H) \\ \hline \mathit{ParticipatesIn}_{\text{PTA}}(s, \Phi_M) \end{array}$$

It is straightforward to show that the conclusion follows from the premises. By the second conjunct of theorem (449.1), the first premise implies that $H \Rightarrow M$, i.e., $\Box\forall x(Hx \rightarrow Mx)$. By the T schema, we may infer $\forall x(Hx \rightarrow Mx)$. By theorem (454.1), the second premise of the argument implies Hs. Hence, it follows that Ms. But, again by theorem (454.1), it now follows that $\mathit{ParticipatesIn}_{\text{PTA}}(s, \Phi_M)$.[264]

(**464**) **Remark:** The Non-Identity Principle. Recall that in the section on the thin conception of the Forms, in Remark (439), we noted Vlastos's claim that the following principle was implicitly presupposed in the Third Man Argument:

(NI) **Non-Identity Principle**
If x is F, x is not identical with The Form of F.

We've already discussed a version of (NI) in connection with the thin conception of Forms. We leave it as an exercise for the reader to explain why, on the thick conception of Forms, object theory not only rejects $Fx \rightarrow x \neq \Phi_F$ but also $xF \rightarrow x \neq \Phi_F$. Instead we consider the corresponding versions of (NI) when the antecedent is formally represented by the two different forms of participation. For then we obtain:

(NIa) If x participates$_{\text{PTA}}$ in The Form of F, x is not identical with that Form.
$\mathit{ParticipatesIn}_{\text{PTA}}(x, \Phi_F) \rightarrow x \neq \Phi_F$

(NIb) If x participates$_{\text{PH}}$ in The Form of F, x is not identical with that Form.
$\mathit{ParticipatesIn}_{\text{PH}}(x, \Phi_F) \rightarrow x \neq \Phi_F$

[264] This discussion corrects an error in Pelletier & Zalta 2000. In that paper, we analyzed the major premise as $\mathit{ParticipatesIn}_{\text{PH}}(\Phi_H, \Phi_M)$, mistakenly believing that this implied that $\Phi_H M$, i.e., that this implied that The Form of Humanity is mortal *pros heauto*. But the discussion in (453) explains why this is an error. The error is corrected in the analysis above: "Humans are mortal" is analyzed as the claim that The Form of Humanity is mortal *pros heauto*. *Given* this as a premise, we may infer via the second conjunct of theorem (449.1) that $H \Rightarrow M$.

It is interesting to consider why *both* claims are rejected by the present theory.

To disprove (NIa), we produce an x and F such that $\mathit{ParticipatesIn}_{\text{PTA}}(x, \Phi_F)$ and $x = \Phi_F$. But let x be $\Phi_{A!}$ and F be $A!$. Then $\mathit{ParticipatesIn}_{\text{PTA}}(\Phi_{A!}, \Phi_{A!})$, by (461.2), and since $\Phi_{A!}\downarrow$, Rule =I yields $\Phi_{A!} = \Phi_{A!}$.

To disprove (NIb), we produce an x and F such that $\mathit{ParticipatesIn}_{\text{PH}}(x, \Phi_F)$ and $x = \Phi_F$. But let our witnesses be Φ_G and G. Then $\mathit{ParticipatesIn}_{\text{PH}}(\Phi_G, \Phi_G)$, by (458), and since $\Phi_G\downarrow$, Rule =I yields $\Phi_G = \Phi_G$.

(**465**) **Remark:** The Third Man Argument Under the Thick Conception. It is now straightforward to see how the Third Man Argument (TMA) is undermined in object theory. We begin with the fact that the existence of two forms of predication led Pelletier & Zalta 2000 to distinguish two forms of participation. Consequently, here are two versions of each of (OM), (SP), and (NI):

- The two versions of (OM), in (456.1) and (456.2), are both provably true.
- The two versions of (SP) fare differently. In (457), we saw that when we interpret the 'is' in "The Form of F is F" as exemplification, the result is subject to counterexample. And in (458), we saw that when we interpret the 'is' in "The Form of F is F" as encodes, the result is provably true. Let us henceforth use the abbreviations (SPa) and (SPb) to represent these two formalized versions of (SP), respectively:

  (SPa) $F\Phi_F$

  (SPb) $\Phi_F F$
- The two versions of (NI), namely (NIa) and (NIb), were discussed in (464), and both are provably false.

It should also be clear that (NIa) and (SPa) are inconsistent, given (454.1) and the reflexivity of identiy.[265]

Similarly, we can distinguish two versions of TMA in light of these results. If we assume for the moment (contrary to what we've established) that all of the principles involved are true, then the original TMA begins with the premise that there are two distinct F-things *pros ta alla* and leads to a contradiction as follows:[266]

> Suppose $Fx \,\&\, Fy \,\&\, x \neq y$. Then by (456.1), it follows that $\exists z(z = \Phi_F \,\&\, \mathit{ParticipatesIn}_{\text{PTA}}(x,z) \,\&\, \mathit{ParticipatesIn}_{\text{PTA}}(y,z))$. Assume that a is such an object, so that we know $a = \Phi_F \,\&\, \mathit{ParticipatesIn}_{\text{PTA}}(x,a) \,\&\, \mathit{ParticipatesIn}_{\text{PTA}}(y,a)$. The first and second conjunct of this result imply $\mathit{ParticipatesIn}_{\text{PTA}}(x, \Phi_F)$, and

[265] From (SPa) it follow by (454.1) that $\mathit{ParticipatesIn}_{\text{PTA}}(\Phi_F, \Phi_F)$, and from this it follows by (NIa) that $\Phi_F \neq \Phi_F$. So (NIa) and (SPa) can't both be true.

[266] The following analysis differs from that of Pelletier & Zalta 2000 in that we now derive the contradiction without appealing to the Uniqueness Principle (i.e., that the Form of F is unique.

so by (NIa) in (464), it follows that $x \neq \Phi_F$. By (SPa) above, we know $F\Phi_F$. Since we now know that Fx, $F\Phi_F$, and $x \neq \Phi_F$, it follows by (456.1) that $\exists z(z = \Phi_F \; \& \; ParticipatesIn_{\text{PTA}}(x,z) \; \& \; ParticipatesIn_{\text{PTA}}(\Phi_F, z))$. Assume that b is such an object, so that we know $b = \Phi_F \; \& \; ParticipatesIn_{\text{PTA}}(x,b) \; \& \; ParticipatesIn_{\text{PTA}}(\Phi_F, b)$. The first and third conjunct of this last result imply $ParticipatesIn_{\text{PTA}}(\Phi_F, \Phi_F)$, and so by (NIa) above, it follows that $\Phi_F \neq \Phi_F$. But this contradicts $\Phi_F = \Phi_F$, which we know by the fact that $\Phi_F\downarrow$ and Rule =I (118.1).

Our discussion thus far has produced a number of reasons why the above argument isn't sound.

Again, assuming for the moment that all the principles are true, the second version of the Third Man Argument begins with the premise that there are two distinct F-things *pros heauto* and leads to a contradiction as follows:[267]

Suppose xF, yF, and $x \neq y$. Then it follows by (456.2) that $\exists z(z = \Phi_F \; \& \; ParticipatesIn_{\text{PH}}(x,z) \; \& \; ParticipatesIn_{\text{PH}}(y,z))$. Assume c is such an object, so that we know:

$$c = \Phi_F \;\; \& \;\; ParticipatesIn_{\text{PH}}(x,c) \;\; \& \;\; ParticipatesIn_{\text{PH}}(y,c)$$

$ParticipatesIn_{\text{PH}}(x, \Phi_F)$, which by (NIb), implies $x \neq \Phi_F$. Moreover, by (458), we know $\Phi_F F$. Hence, we know xF, $\Phi_F F$, and $x \neq \Phi_F$. So by (456.2), it follows that $\exists z(z = \Phi_F \; \& \; ParticipatesIn_{\text{PH}}(x,z) \; \& \; ParticipatesIn_{\text{PH}}(\Phi_F, z))$. Assume d is such an object, so that we know:

$$d = \Phi_F \;\; \& \;\; ParticipatesIn_{\text{PH}}(x,d) \;\; \& \;\; ParticipatesIn_{\text{PH}}(\Phi_F, d)$$

The first and third conjunct of this result imply $ParticipatesIn_{\text{PH}}(\Phi_F, \Phi_F)$, which by (NIb) implies $\Phi_F \neq \Phi_F$. But this contradicts $\Phi_F = \Phi_F$, which we know by the fact that $\Phi_F\downarrow$ and Rule =I.

In this case, the soundness of the argument is undermined by the failure of (NIb), as described at the end of (464), since the version of the Self-Predication Principle used in the argument is a theorem (458). These considerations put the Third Man Argument to rest, under both the thick and thin conceptions of the Forms.

[267] See Pelletier & Zalta 2000 (174), and Frances 1996 (59).

Chapter 12

Situations, Worlds, Times, and Stories

In this chapter we develop the theory of situations, worlds (both possible and impossible), moments of time, and world-states. Along the way we introduce three kinds of world-indexed objects: world-indexed truth-values, classes and relations. In what follows, we do not distinguish propositions and states of affairs. Indeed, we shall often refer to 0-ary relations as *states of affairs*, since that more closely follows the traditional language of situation theory.

12.1 Situations

(**466**) **Remark:** On the Nature of Situations. In a series of papers (1980, 1981a, 1981b) that culminated in their book of 1983, Barwise and Perry argued against views that were widely held in the field of natural language semantics. Here are some of the widely held views that they criticized: (a) that possible worlds, taken as primitive, constitute a fundamental semantic domain for interpreting natural language, (b) that properties (and relations) are analyzable as functions from possible worlds to sets of (sequences of) individuals, (c) that the denotation of a sentence is a truth-value, and (d) that the denotation of a sentence shifts when the sentence appears in indirect, intensional contexts. Barwise and Perry suggested that a better semantic theory of language could be developed if possible worlds were replaced with *situations*, i.e., parts of the world in which one or more states of affairs hold, where states of affairs consist of objects standing in relations.

Barwise and Perry (1984, 23) subsequently realized that their book of 1983 offered a *model* of situations rather than a *theory* of them and, consequently, changed the direction of their research. In their early attempts to develop a theory of situations, they brought to bear certain intuitions they had previously had about the nature of situations. The fundamental intuition never wavered (Barwise 1985, 185):

> By a situation, then, we mean a part of reality that can be comprehended as a whole in its own right—one that interacts with other things. By interacting with other things, we mean that they have properties or relate to other things. They can be causes and effects, for example, as when we see them or bring them about. Events are situations, but so are more static situations, even eternal situations involving mathematical objects. We use $s, s', s'', \ldots$ to range over real situations. There is a binary relation $s \models \sigma$, read "σ holds in s", that holds between various situations s and states of affairs σ; that is, situations and states of affairs are the appropriate arguments for this relation of holding in.

One of the guiding intuitions was the distinction between the internal and external properties of situations. In Barwise and Perry 1981b (388), we find:

> Situations have properties of two sorts, internal and external. The cat's walking on the piano distressed Henry. Its doing so is what we call an external property of the event. The event consists of a certain cat performing a certain activity on a certain piano; these are its internal properties.

This distinction between the internal and external properties appears throughout the course of publications on situation theory. For example, Barwise writes (1985, 185):

> If $s \models \sigma$, then the fact σ is called a fact of s, or more explicitly, a fact about the internal structure of s. There are also other kinds of facts about s, facts external to s, so the difference between being a fact that holds in s and a fact about s more generally must be borne in mind.

And in Barwise 1989a (263–4), we find:

> The facts determined by a particular situation are, at least intuitively, intrinsic to that situation. By contrast, the information a situation carries depends not just on the facts determined by that situation but is relative to constraints linking those facts to other facts, facts that obtain in virtue of other situations. Thus, information carried is not usually (if ever) intrinsic to the situation.
>
> The objects which actual situations make factual thus play a key role in the theory. They serve to characterize the intrinsic nature of a situation.

Interestingly, the intuitive distinction between the intrinsic, internal properties of a situation and its extrinsic, external properties, never made it to the level of theory; situation theorists never formally regimented the distinction.

However, we take it to be the key to the analysis of situations in what follows. If we identify a situation s to be an abstract object that encodes only properties of the form $[\lambda y\, p]$ (where p is some state of affairs or proposition), so that we can say p is *encoded in* s whenever s encodes $[\lambda y\, p]$, then the intrinsic, internal properties of s are its encoded properties, while the extrinsic, external properties of s will be any exemplification facts about s or any logical or natural-law-based exemplification generalizations either about s or about the relations in the states of affairs encoded in s.

This analysis of situations was worked out originally in Zalta 1993. In what follows, we reprise the most important definitions and theorems from that paper, as well as many new ones. Readers familiar with the classical works of situation theory in the Barwise & Perry tradition should note that the *infons* of later situation theory:

$$\langle\langle R^n, a_1, \ldots, a_n; 1\rangle\rangle$$

$$\langle\langle R^n, a_1, \ldots, a_n; 0\rangle\rangle$$

are simply states of affairs in which objects $a_1, \ldots, a_n$ do or do not stand in relation R^n, depending on whether the polarity is 1 or 0. Thus, we have no need of infon notation, since we not only have the standard notation $R^n a_1 \ldots a_n$ and $\neg R^n a_1 \ldots a_n$ but also the λ-notation $[\lambda\, R^n a_1 \ldots a_n]$ and $[\lambda\, \neg R^n a_1 \ldots a_n]$, both of which denote states of affairs. Moreover, we shall represent the following situation-theoretic claims asserting that an infon holds in situation s:

$$s \models \langle\langle R^n, a_1, \ldots, a_n; 1\rangle\rangle$$

$$s \models \langle\langle R^n, a_1, \ldots, a_n; 0\rangle\rangle$$

more simply as follows:

$$s \models R^n a_1 \ldots a_n$$

$$s \models \neg R^n a_1 \ldots a_n$$

These claims will be defined, respectively, as:

$$s[\lambda y\, R^n a_1 \ldots a_n]$$

$$s[\lambda y\, \neg R^n a_1 \ldots a_n]$$

We shall, on occasion, point out (a) where axioms central to situation theory are derived as theorems, and (b) where unsettled questions of situation theory, as collated in Barwise 1989a, are settled by a theorem of object theory. Note that our work thus far has already resolved Choices 14–17 in Barwise 1989a (270–1). The Comprehension Principle for Propositions (i.e., states of affairs), derived as theorem (194), guarantees that we can freely form states of affairs out of any objects and relations (Choice 14); that not every state of affairs is basic (Choice 15, Alternative 15.2); that there is a rich algebraic structure on the space of states of affairs (Choice 16); and that every state of affairs has a dual (Choice 17). Moreover, theorem (268.3) resolves Choice 13 in favor of Alternative 13.2: for each property F, there are *distinct* abstract objects a, b such that $[\lambda Fa] = [\lambda Fb]$. In such cases, the states of affairs $[\lambda Fa]$ and $[\lambda Fb]$ are identical, but it is not the case that $a = b$. In Barwise 1989a (270), Alternative 13.2 is in effect when $\langle\langle R, a; i\rangle\rangle = \langle\langle S, b; j\rangle\rangle$ fails to imply $R = S$, $a = b$, and $i = j$.

12.1.1 Basic Definitions and Theorems

(**467**) **Definition and Theorems:** Situations. Using the notion of propositional property defined in (275), we may say that x *is a situation* just in case x is an abstract object that encodes only propositional properties:

(.1) $\mathit{Situation}(x) \equiv_{df} A!x \;\&\; \forall F(xF \rightarrow \mathit{Propositional}(F))$

By the definition (275), it follows that a situation is an abstract object x such that every property F that x encodes is a property of the form $[\lambda y\, p]$, for some proposition p. Note that the above definition of situation decides Choice 9 in Barwise 1989a (267) since it is easy to show that there are objects (e.g., ordinary objects and abstract objects that encode non-propositional properties) that are not situations.

It is easy to establish that (.2) situations exist:

(.2) $\exists x \mathit{Situation}(x)$

Let R be any relation, and a and b be any objects, and consider the instance of the Comprehension Principle for Abstract Objects (53) that asserts $\exists x(A!x \;\& \;\forall F(xF \equiv F = [\lambda y\, Rab] \vee F = [\lambda y\, \neg Rba]))$. Let c be such an object. Clearly, it is provable that c encodes just the two properties $[\lambda y\, Rab]$ and $[\lambda y\, \neg Rba]$. So every property c encodes is a propositional property. Hence, c is a situation; intuitively, c is the 'smallest' situation that encodes $[\lambda y\, Rab]$ and $[\lambda y\, \neg Rba]$.

Note that (.3) $\mathit{Situation}(\kappa)$ implies $\kappa\downarrow$, for any individual term κ. That is, it is a modally strict theorem schema, for any individual term κ, that:

(.3) $\mathit{Situation}(\kappa) \rightarrow \kappa\downarrow$

Given (.2) and (.3), it follows from the metadefinition in (336) that $\mathit{Situation}(x)$ is a restriction condition: it has a single free variable, it is $\Box$-theorem that $\exists x \mathit{Situation}(x)$ (i.e., it is a strictly non-empty condition), and it is a $\Box$-theorem that $\mathit{Situation}(\kappa) \rightarrow \kappa\downarrow$ (i.e., it has strict existential import). But before we introduce restricted variables, we establish a theorem needed to show that $\mathit{Situation}(x)$ is a rigid restriction condition.

(**468**) **Theorem:** Some Known Situations. Given our definitions, it follows that truth-values are situations:

$$TruthValue(x) \to Situation(x)$$

From this and the theorems that $TruthValue(\circ p)$ (301)★, $TruthValue(\top)$ (303.1)★, and $TruthValue(\bot)$ (303.2)★, we can infer, as ★-theorems, that $\circ p$, $\top$, and $\bot$ are all situations. But see the discussion following the next group of lemmas, which explain why we can conclude that in fact $Situation(\circ p)$, $Situation(\top)$, and $Situation(\bot)$ are all modally strict theorems.

(**469**) **Lemmas:** Modal Collapse of $Situation(x)$ and Restricted Variables for Situations. The following lemmas prove useful. (.1) x is a situation if and only if necessarily x is a situation; (.2) possibly x is a situation if and only if x is a situation; (.3) possibly x is a situation if and only if necessarily x is a situation; and (.4) actually x is a situation if and only if x is a situation:

(.1) $Situation(x) \equiv \Box Situation(x)$

(.2) $\Diamond Situation(x) \equiv Situation(x)$

(.3) $\Diamond Situation(x) \equiv \Box Situation(x)$

(.4) $\mathcal{A}Situation(x) \equiv Situation(x)$

Note that by applying RN to the left-to-right direction of (.1), it follows that $\Box(Situation(x) \to \Box Situation(x))$ is a theorem. In other words, $Situation(x)$ is modally collapsed. Since we established, at the end of (468), that $Situation(\circ p)$, $Situation(\top)$, and $Situation(\bot)$ are all theorems, it follows by the Rule of Modal Strictness (173) that these latter are all modally strict theorems:

(.5) $Situation(\circ p)$

(.6) $Situation(\top)$

(.7) $Situation(\bot)$

Note further that by applying GEN to the left-to-right direction of (.1), it we know that $\vdash_\Box \forall x(Situation(x) \to \Box Situation(x))$. Hence $Situation(x)$ is a rigid condition on objects, as defined in (260.1).

Thus, $Situation(x)$ not only meets the definition of a restriction condition, as we saw in (467), but also that of a rigid restriction condition, as defined in (340). So, we may henceforth use the symbols $s, s', s'', \ldots$ as rigid restricted variables ranging over situations and $s_1, s_2, s_3, \ldots$ as constants denoting situations. And, given our discussion of reasoning with restricted variables in (340) and (341), we can both (a) assert theorems and instantiate modal principles with free situation variables, and (b) reason with bound and free occurrences of these variables. See Remark (472) below.

From the fact that there are situations, we also know that the quantifiers binding the restricted variable s, $\forall s$ and $\exists s$, behave classically in the sense that $\forall s\varphi \to \exists s\varphi$ is a theorem; cf. Remark (342). Moreover, from the fact that $Situation(\kappa) \to \kappa\downarrow$, for any individual term κ, we may use the variable s in the definiens and definiendum of a definition-by-$\equiv$ without having to add existence clauses to the definiens; for any instance of the definition in which a provably empty term has been uniformly substituted for s, the negation of the definiendum is provable.

12.1.2 Truth in a Situation

(**470**) **Definition:** Truth In a Situation. Recall that in (295) we defined: x encodes p, written $x\Sigma p$, just in case x exists and encodes $[\lambda y\, p]$. We now say that proposition p is *true in* x, written $x \models p$, just in case x is a situation that encodes p:

$$x \models p \ \equiv_{df} \ Situation(x) \,\&\, x\Sigma p$$

Given our conventions in (338.1), we may use restricted variables to recast the above definition by saying that p is *true in* s just in case s encodes p:

$$s \models p \ \equiv_{df} \ s\Sigma p$$

Note that there are other ways to read this definition. If we regard 0-ary relations as *states of affairs*, then the variable p ranges over such states and we may then read '$s \models p$' in situation-theoretic terms as follows:

State of affairs p *holds in* situation s

State of affairs p *is a fact in* situation s

Situation s *makes* p *true*

Finally, note that in what follows, we always read '$\models$' with smallest possible scope. So, for example, $x \models p \to p$ is to be parsed as $(x \models p) \to p$ rather than $x \models (p \to p)$ and $s \models q \,\&\, \neg q$ is to be parsed as $(s \models q) \,\&\, \neg q$.

(**471**) **Lemma:** Truth in a Situation and Encoding. Whenever x is a situation, then p is true in x if and only if x encodes *being such that* p:

$$Situation(x) \to ((x \models p) \equiv x[\lambda y\, p])$$

Using restricted variables, this can be expressed as $(s \models p) \equiv s[\lambda y\, p]$.

(**472**) **Remark:** Rule RN and Restricted Variables. Readers who skipped the discussion of reasoning with restricted variables in (340) and (341) may find the following discussion useful. It is extremely important to remember that when reasoning with rigid restricted variables, we are making use of the extended rules RN and RA justified in (341.3.a) and (341.3.b), respectively. For example, as we just saw, we can express (471) using restricted variables as:

(A) $(s \models p) \equiv s[\lambda y\, p]$

But whereas (B) does follow from (A), it doesn't follow by the Rule RN formulated in (68), but by the expanded version in (341):

(B) $\Box((s \models p) \equiv s[\lambda y\, p])$

Though (B) is in fact a theorem, it isn't provable by a single application of RN to (A). To see why this inference isn't valid, consider first that when we eliminate the restricted variable, then (B) is given a conditional interpretation and so asserts:

(C) $Situation(x) \to \Box((x \models p) \equiv x[\lambda y\, p])$

But, clearly, (C) is not derivable from (471) by Rule RN (68). However, as we noted, (B), i.e., (C), is a theorem:

> *Proof*. Assume $Situation(x)$. So by (469.1), $\Box Situation(x)$. Moreover, since (471) is a modally strict theorem, we can apply Rule RM (157.1) to conclude:
>
> $$\Box Situation(x) \to \Box((x \models p) \equiv x[\lambda y\, p])$$
>
> Hence $\Box((x \models p) \equiv x[\lambda y\, p])$. So by conditional proof, (C), i.e., (B).

The key fact about this reasoning is that in order to derive (B), we had to appeal to the left-to-right direction of theorem (469.1), which tells us that if x is a situation, then necessarily x is a situation. Our expanded rule RN, proved in (341.3.a), tells us that when we have a *rigid* restricted variable, then we can assume that it ranges over objects that *necessarily* satisfy the restriction condition on the variable whenever they satisfy the condition. For further discussion, see Remarks (337), (340), and (341).

(**473**) **Lemmas:** Rigidity of Truth In a Situation. Our definitions and theorems also guarantee that: (.1) p is true in s if and only if it is necessary that p is true in s; (.2) it is possible that p is true in s if and only if p is true in s; (.3) it is possible that p is true in s if and only if it is necessary that p is true in s; (.4) actually p is true in s if and only if p is true in s; and (.5) p fails to be true in s if and only if necessarily p fails to be true in s:

(.1) $s \models p \equiv \Box s \models p$

(.2) $\Diamond s \models p \equiv s \models p$

(.3) $\Diamond s \models p \equiv \Box s \models p$

(.4) $\mathcal{A}s \models p \equiv s \models p$

(.5) $\neg s \models p \equiv \Box \neg s \models p$

The proof of (.1) in the Appendix is in two parts. First, we prove the claim by eliminating the restricted variable. Then we show how to prove the claim as it is stated above with restricted variables, by making use of what we learned in Remark (472) and in our discussion of reasoning with restricted variables in (340) and (341).

12.1.3 Situation Identity and Parts of Situations

(**474**) **Theorem:** Situation Identity. A fundamental fact concerning situation identity is now derivable, namely, that situations s and s' are identical if and only if they make the same propositions true:

$$s = s' \equiv \forall p(s \models p \equiv s' \models p)$$

This decides Choice 5 in Barwise 1989a (264) in favor of Alternative 5.1.

(**475**) **Definition:** Parts of Situations. We say that situation *s is a part of* situation s', written $s \trianglelefteq s'$, just in case every proposition true in s is true in s':

$$s \trianglelefteq s' \ \equiv_{df} \ \forall p(s \models p \rightarrow s' \models p)$$

This definition determines Choice 2 in Barwise 1989a (261), since it requires that every part of a situation be a situation. By our conventions for using free restricted variables in definitions-by-$\equiv$ (338.2), the above definition is short for:

$$x \trianglelefteq y \ \equiv_{df} \ Situation(x) \,\&\, Situation(y) \,\&\, \forall p(x \models p \rightarrow y \models p)$$

(**476**) **Theorems:** *Part of* is a Partial Order on Situations. *Part of* is reflexive, anti-symmetric, and transitive on the situations:

(.1) $s \trianglelefteq s$

(.2) $s \trianglelefteq s' \,\&\, s \neq s' \rightarrow \neg(s' \trianglelefteq s)$

(.3) $s \trianglelefteq s' \,\&\, s' \trianglelefteq s'' \rightarrow s \trianglelefteq s''$

Whereas a partial ordering of situations by $\trianglelefteq$ is assumed in situation theory (cf. Barwise 1989b 185; 1989a, 259), such an ordering falls out as a theorem of object theory.

(**477**) **Theorems:** Parts and Identity. Two other constraints on the identity of situations are derivable, namely, that (.1) situations s and s' are identical if and only if each is part of the other, and that (.2) situations s and s' are identical if and only if they have the same parts:

(.1) $s = s' \equiv s \trianglelefteq s' \,\&\, s' \trianglelefteq s$

(.2) $s = s' \equiv \forall s''(s'' \trianglelefteq s \equiv s'' \trianglelefteq s')$

In the usual way, when $s \trianglelefteq s'$ and $s \neq s'$, it may be useful to say that s is a *proper* part of s'.

(**478**) **Definition:** Persistency. In situation theory, a state of affairs p is *persistent* if and only if whenever p holds in a situation s, p holds in every situation s' of which s is a part:

$$Persistent(p) \equiv_{df} \forall s(s \models p` \rightarrow \forall s'(s \trianglelefteq s' \rightarrow s' \models p))$$

Cf. Barwise 1989a (265).

(**479**) **Theorem:** Propositions are Persistent. The following is therefore an immediate consequence of the definitions of $\trianglelefteq$ and *Persistent*:

$$\forall p\, Persistent(p)$$

Thus, our theory implies Alternative 6.1 at Choice 6 in Barwise 1989a (265).[268]

12.1.4 Comprehension Conditions for Situations

(**480**) **Metadefinition:** Conditions on Propositional Properties. In what follows, let us say that φ is a *condition on propositional properties* if and only if there is a modally strict proof of: every property such that φ is a propositional property, i.e.,

φ is a *condition on propositional properties* if and only if
$\vdash_\Box \forall F(\varphi \rightarrow Propositional(F))$

Note that, on this definition, $F \neq F$ counts as a condition on propositional properties; one can derive the definiens by failure of the antecedent and GEN. But no propositional property satisfies $F \neq F$. However, the definition rules out formulas that are satisfied by non-propositional properties.[269]

[268]By its very set-up, our theory assumes Alternative 11.1 of Choice 11 (Barwise 1989a, 268): no relations are perspectival. The argument places of a relation are now allowed to vary and so if a relation R is a constituent of a state of affairs that is true in situation s, then R has the same argument places in any situation s' of which s is a part.

[269]For example, let ψ be the condition $F = F \,\&\, \neg Pa$, where Pa is some contingently true fact asserted as a modally fragile axiom. Intuitively, ψ shouldn't count as a condition on propositional

(**481**) **Theorems:** An Important Equivalence. Whenever φ is a condition on propositional properties, then x is a situation that encodes all and only the properties F such that φ if, and only if, x is an abstract object that encodes all and only the properties F such that φ:

$(Situation(x)\,\&\,\forall F(xF \equiv \varphi)) \equiv (A!x\,\&\,\forall F(xF \equiv \varphi))$,
provided φ is a condition on propositional properties.

This theorem makes it easier to derive comprehension conditions for situations.

(**482**) **Theorems:** Comprehension Conditions for Situations. Where φ is a condition on propositional properties in which x doesn't occur free, there exists a (unique) situation x that encodes all and only those properties F such that φ:

(.1) $\exists x(Situation(x)\,\&\,\forall F(xF \equiv \varphi))$, provided φ is a condition on propositional properties in which x doesn't occur free

(.2) $\exists!x(Situation(x)\,\&\,\forall F(xF \equiv \varphi))$, provided φ is a condition on propositional properties in which x doesn't occur free

Thus, whereas the classic works in situation theory (e.g., Barwise and Perry 1983, 7–8; Barwise 1989a, 261) assume the existence of situations, object theory yields, as theorems, existence principles that comprehend the domain of situations.

Notice also that with these theorems, we have developed a precise *theory* of situations, since (.1) and (474), respectively, constitute fully general comprehension and identity principles for situations.

(**483**) **Theorems:** Canonical Situation Descriptions. It now follows that when φ is a condition on propositional properties, there exists something which is *the* situation that encodes exactly the properties F such that φ:

(.1) $\imath x(Situation(x)\,\&\,\forall F(xF \equiv \varphi))\downarrow$, provided φ is a condition on propositional properties in which x doesn't occur free

Moreover, it is straightforward to show:

(.2) $\imath x(Situation(x)\,\&\,\forall F(xF \equiv \varphi)) = \imath x(A!x\,\&\,\forall F(xF \equiv \varphi))$, provided φ is a condition on propositional properties in which x doesn't occur free

---

properties: at worlds where Pa is false, *every* property is such that ψ. This intuition is preserved, for we can't show that ψ is a condition on propositional properties. Since Pa is a (contingent) axiom, it follows that $\neg(F = F\,\&\,\neg Pa)$, i.e., $\neg\psi$. So $\psi \rightarrow Propositional(F)$, by (77.3). By GEN, $\forall F(\psi \rightarrow Propositional(F))$. But this universal generalization was proved by appeal to a contingency, namely Pa, and so fails to be a modally strict theorem, as required by the definition.

Hence, when φ is a condition on propositional properties, we may say that descriptions of the form $\imath x(Situation(x) \,\&\, \forall F(xF \equiv \varphi))$ and $\imath s\forall F(sF \equiv \varphi)$ are *canonical situation descriptions*. If permitted a modest abuse of language, we may say that these canonical descriptions denote *canonical situations* (cf. (253)), though see (486) for an simpler, alternative method of describing canonical situations.

(**484**) **Theorems:** Canonical Situations and Rigid Conditions on Properties. Since we've defined what it is for φ to be a condition on propositional properties (480) and what it is for φ to be a rigid condition on properties (260.1), we may combine the two to talk about formulas φ that are *rigid conditions on propositional properties*. Consequently, whenever φ is a rigid condition on propositional properties, then it is a modally strict fact that if something is identical to a canonical situation, then it encodes exactly the properties such that φ:

> $y = \imath s\forall F(sF \equiv \varphi) \rightarrow \forall F(yF \equiv \varphi)$, provided φ is a rigid condition on propositional properties in which x isn't free.

(**485**) **Metadefinitions:** Strict Canonicity and Situations. Given the preceding results, we say that $\imath s\forall F(sF \equiv \varphi)$ is a *strictly canonical situation description* whenever φ is a rigid condition on propositional properties. In the usual way, we sometimes abuse language to say that $\imath s\forall F(sF \equiv \varphi)$ is a *strictly canonical situation*. Though this notion has now been defined in a familiar way, (486) below offers simpler comprehension conditions for situations and simpler canonical and strictly canonical descriptions for situations. Before we examine these developments, the reader may wish to try the following.

Exercises:

1. Show that $\circ p$ (294), $\top$ (302.1), and $\bot$ (302.2) are (identical to) canonical situations.

2. Consider the following situation description:

   $$\imath s\forall F(sF \equiv F = [\lambda y\, Rab] \vee F = [\lambda y\, \neg Rba])$$

   (a) Show that this is a canonical situation description; i.e., show that the condition $F = [\lambda y\, Rab] \vee F = [\lambda y\, \neg Rba]$ is a condition on propositional properties. (b) Show that this is a strictly canonical situation description.

3. Show that $\circ p$, $\top$, and $\bot$ are not (identical to) strictly canonical situations.

(**486**) **Theorems:** Simplified Comprehension Conditions for Situations and Simpler Notions of (Strict) Canonicity. Without too much effort, one can show that when φ is any formula in which s doesn't occur free, there exists a situation that makes true all and only the propositions such that φ:

(.1) $\exists s\forall p(s \models p \equiv \varphi)$, provided s doesn't occur free in φ

Clearly, *every* formula φ constitutes a condition on propositions – if the variable s occurs free in φ, then choose a variable, say s', that doesn't occur free in φ and then the alphabetic variant $\exists s'\forall p(s' \models p \equiv \varphi)$ becomes derivable. So the above theorem schema offers unrestricted comprehension conditions for situations. This schema proves its worth primarily when one considers instances in which the variable p occurs free in φ. But if the variable p doesn't occur free in φ, then φ places a vacuous condition on propositions – if φ is true, then every proposition satisfies such a φ (with the resulting instance asserting the existence of the a situation that encodes every proposition) and if φ is false, then no proposition does (with the resulting instance asserting the existence of a situation that encodes no propositions).

It follows, in the usual way, that (.2) there exists a unique situation that makes true all and only the propositions p such that φ, and that (.2) the situation that makes true all and only the propositions p such that φ exists:

(.2) $\exists! s\forall p(s \models p \equiv \varphi)$, provided s doesn't occur free in φ

(.3) $\imath s\forall p(s \models p \equiv \varphi)\downarrow$, provided s doesn't occur free in φ

Thus, we now have another canonical way of describing situations, namely, in terms of the descriptions in (.3). So let us we continue with our innocuous abuse of language and say that $\imath s\forall p(s \models p \equiv \varphi)$ is a *canonical situation*. Moreover, if φ is a rigid condition on propositions, i.e., if $\vdash_\Box \forall p(\varphi \rightarrow \Box\varphi)$, then $\imath s\forall p(s \models p \equiv \varphi)$ is a *strictly* canonical situation. Thus we have the following analogue of (484):

(.4) $y = \imath s\forall p(s \models p \equiv \varphi) \rightarrow \forall p(y \models p \equiv \varphi)$,
provided φ is a rigid condition on propositions

Exercise: In light of the fact that truth-values are situations (468), we can simplify the definition of $TruthValueOf(x,p)$ that we introduced in Section 10.1. Instead of the definition given in (286):

$$TruthValueOf(x,p) \equiv_{df} A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])) \qquad (286)$$

we may instead define:

$$TruthValueOf(s,p) \equiv_{df} \forall q(s \models q \equiv (q \equiv p))$$

In other words, a situation s is a truth-value of p just in case all and only the propositions q that are materially equivalent to p are true in s. Show that this new definition is equivalent to (286). Then simplify and definitions and proofs of the theorems in Section 10.1.[270] Then show how to revise the definition of φ-*AbstractionOf*(x,p) in (385.1), and derive (386.1), (386.3), and (387.1) in light of the new definition.

[270]Since the revised definition of $TruthValueOf(s,p)$ appears to be simpler than the original in

12.1.5 Null and Trivial Situations

(**487**) **Definitions:** Null and Trivial Situations. We define: (.1) x is a *null situation* if and only if x is a situation in which no propositions are true, and (.2) x is a *trivial situation* iff x is a situation in which every proposition is true:

(.1) $NullSituation(x) \equiv_{df} Situation(x) \,\&\, \neg\exists p(x \models p)$

(.2) $TrivialSituation(x) \equiv_{df} Situation(x) \,\&\, \forall p(x \models p)$

(**488**) **Theorems:** Existence and Uniqueness of Null and Trivial Situations. It is now easily established that (.1) there is a unique null situation and that (.2) there is a unique trivial situation:

(.1) $\exists! x NullSituation(x)$

(.2) $\exists! x TrivialSituation(x)$

Consequently, it follows, by a modally strict proof that (.3) the null situation exists, and that (.4) the trivial situation exists:

(.3) $\imath x NullSituation(x)\downarrow$

(.4) $\imath x TrivialSituation(x)\downarrow$

(**489**) **Definitions:** Notation for The Null Situation and The Trivial Situation. We now introduce:

(.1) $\boldsymbol{s}_{\varnothing} =_{df} \imath x NullSituation(x)$

(.2) $\boldsymbol{s}_{\boldsymbol{V}} =_{df} \imath x TrivialSituation(x)$

In the notation $\boldsymbol{s}_{\varnothing}$ and $\boldsymbol{s}_{\boldsymbol{V}}$, we use a boldface, italic $\boldsymbol{s}$ as part of the name. This boldface, italic $\boldsymbol{s}$ is to be distinguished from the (nonbold, italic) restricted variable s. The reason for this should be clear: expressions such as $s_{\varnothing}$ and s_V

---

(286), one might wonder why I didn't change the order of presentation, and expound the theory of situations prior to the theory of truth-values. I did consider revising the present work so as to re-order the presentation, but in the end, I decided against doing so, for two reasons: (1) Historically, *truth-values* were systematized first; they are the very first logical objects introduced in Frege's *Grundgesetze* (1893) and they are introduced right at the outset, in I.1.§2, right after the introduction of functions in I.1.§1. Given the subsequent importance of truth-values in the development of contemporary logic and the relatively recent systemization of *situations* in Barwise & Perry 1983, it seemed appropriate to present the theory of truth-values before presenting the theory of situations. (2) There is a extended network of defined notions and theorems that have to be introduced before one can prove the Simplified Comprehension Conditions for Situations in (.1) above. It seemed to me that one could, and should, introduce truth-values directly, without all of the intermediary notions of *situation*, *truth in a situation*, *conditions on propositional properties*, theorems (481), (482.1), etc., all of which are needed to introduce and prove (.1) above. None of this machinery is strictly necessary to define, and prove facts about, truth-values.

(with nonbold italic s) should be used only when we are introducing functional terms having denotations that vary with the value of s. By contrast, we are here introducing new names for distinguished objects.

(**490**) **Theorems:** Facts About $\boldsymbol{s}_{\varnothing}$ and $\boldsymbol{s}_{\boldsymbol{V}}$. It is now to be established that (.1) if x is a null situation x is necessarily so, and (.2) if x is a trivial situation, x is necessarily so:

(.1) $NullSituation(x) \rightarrow \Box NullSituation(x)$

(.2) $TrivialSituation(x) \rightarrow \Box TrivialSituation(x)$

From these, we may produce modally strict proofs of the following:

(.3) $NullSituation(\boldsymbol{s}_{\varnothing})$

(.4) $TrivialSituation(\boldsymbol{s}_{\boldsymbol{V}})$

(**491**) **Theorems:** Further Facts About Null and Trivial Situations. Recall that we defined the null object $\boldsymbol{a}_{\varnothing}$ in (265.1) as $\imath x Null(x)$, where $Null(x)$ was defined in (263.1) as $A!x \,\&\, \neg\exists F xF$. Recall that we also defined the universal object $\boldsymbol{a}_{\boldsymbol{V}}$ in (265.2) as $\imath x Universal(x)$, where $Universal(x)$ was defined in (263.2) as $A!x \& \forall F xF$. It then follows that: (.1) $NullSituation(x)$ and $Null(x)$ are equivalent conditions; (.2) the null situation is identical to the null object; and (.3) the trivial situation is not identical to the universal object:

(.1) $NullSituation(x) \equiv Null(x)$

(.2) $\boldsymbol{s}_{\varnothing} = \boldsymbol{a}_{\varnothing}$

(.3) $\boldsymbol{s}_{\boldsymbol{V}} \neq \boldsymbol{a}_{\boldsymbol{V}}$

The proof of (.3) is especially interesting, since to show that the trivial situation and the universal object fail to be identical, we must find a property that one encodes which the other doesn't. Since the universal object encodes every property whatsoever and the trivial situation encodes all and only propositional properties, we know the former encodes every property the latter encodes. So one has to show there is a property that the universal object encodes that the trivial situation doesn't. It suffices to prove the existence of a property that is not a propositional property, and the proof of (.3) is interesting because it does just that.

12.1.6 Actual Situations

(**492**) **Definition:** Actual Situations. We say that a situation s *is actual* iff every proposition true in s is true:

$$Actual(s) \equiv_{df} \forall p(s \models p \rightarrow p)$$

Note here that an *actual* situation has been defined in terms of truth, without using the actuality operator. See Remark (494) below.

(**493**) **Theorem:** Some Actual Situations Might Not Be Actual. It is a consequence of our definition of *Actual* that some actual situations might fail to be actual:

$$\exists s(Actual(s) \,\&\, \Diamond\neg Actual(s))$$

The proof of this theorem identifies a witness (i.e., a situation) in which a contingent truth is true. Of course, if s is a situation that is actual because every proposition true in s is a necessary truth, then s couldn't fail to be an actual sitation.

(**494**) **Remark:** Actual vs. Actual* Situations. In (492), we defined an actual situation s to be one such that every proposition true in s is true. We did *not* use an actuality operator in the definiens. This is partly a gesture to the tradition: the actuality of a situation hasn't traditionally been defined in terms of an actuality operator.

But one could say that a situation s is *actual** just in case every proposition true in s is actually true:

$$Actual^*(s) \equiv_{df} \forall p(s \models p \rightarrow \mathcal{A}p)$$

Given such a definition, a theorem emerges that stands in contrast to the fact that actual situations that might not be actual (493). For it is provable that every *actual** situation is necessarily *actual**:

$$Actual^*(s) \rightarrow \Box Actual^*(s)$$

The proof is left to a footnote.[271] In what follows, we shall work primarily with the notion of an *actual* situation. It is important to keep in mind how it differs from the notion of an *actual** situation. In particular, we shall use the notion of an *actual* situation to define an *actual world*, and so some of the theorems governing this notion may not be subject to modally strict proofs.

(**495**) **Theorems:** Actual and Nonactual Situations. It is an immediate consequence that there are both actual and nonactual situations:

[271] Assume $Actual^*(s)$, i.e.,

(ϑ) $\forall p(s \models p \rightarrow \mathcal{A}p)$

We have to show $\Box\forall p(s \models p \rightarrow \mathcal{A}p)$. By BF (167.1), it suffices to show $\forall p\Box(s \models p \rightarrow \mathcal{A}p)$. By GEN, it suffices to show $\Box(s \models p \rightarrow \mathcal{A}p)$, i.e., $\Box(\neg s \models p \vee \mathcal{A}p)$. By (158.15), it suffices to show:

(ξ) $\Box\neg s \models p \vee \Box\mathcal{A}p$

But by (ϑ), we know $s \models p \rightarrow \mathcal{A}p$, i.e., $\neg s \models p \vee \mathcal{A}p$. So we can show ($\xi$) by disjunctive syllogism from this last fact. If $\neg s \models p$, then it follows by (473.5), $\Box\neg s \models p$. Hence (ξ), by $\vee$I. If $\mathcal{A}p$, then $\Box\mathcal{A}p$, by (46.1). Hence (ξ), by $\vee$I.

(.1) $\exists s Actual(s)$

(.2) $\exists s \neg Actual(s)$

The theory therefore decides Choice 4 in Barwise 1989a (262) in favor of Alternative 4.2. Note, moreover, that some propositions are not true in any actual situation:

(.3) $\exists p \forall s(Actual(s) \rightarrow \neg s \models p)$

(**496**) **Lemma:** Embedding Situations in Another Situation. Where s' and s'' are any situations, there exists a situation s of which both s' and s'' are a part and which is a part of every situation s''' of which both s' and s'' are a part:

$$\exists s(s' \trianglelefteq s \ \& \ s'' \trianglelefteq s \ \& \ \forall s'''(s' \trianglelefteq s''' \ \& \ s'' \trianglelefteq s''' \rightarrow s \trianglelefteq s'''))$$

It is tempting to call any such s a *least upper bound* on s' and s'', but we are going to postpone further discussion of algebraic notions until Exercise (505) below.

(**497**) **Theorems:** Facts About Actual Situations. (.1) If p is true in an actual situation s, then s exemplifies being such that p, and (.2) for any two actual situations, there exists an actual situation of which both are a part:

(.1) $Actual(s) \rightarrow (s \models p \rightarrow [\lambda y\ p]s)$

(.2) $(Actual(s') \ \& \ Actual(s'')) \rightarrow \exists s(Actual(s) \ \& \ s' \trianglelefteq s \ \& \ s'' \trianglelefteq s)$

(.2) establishes what Barwise (1989b, 235) calls the Compatibility Principle.

12.1.7 Consistent, Possible and Incompatible Situations

(**498**) **Definition:** Consistency. A situation s is *consistent* iff there is no proposition p such that both p is true in s and $\neg p$ is true in s:

$$Consistent(s) \ \equiv_{df} \ \neg\exists p(s \models p \ \& \ s \models \neg p)$$

(**499**) **Remark:** Types of Consistency. We take the preceding definition of consistency to be a standard one.[272] However, it is interesting to consider how the above definition differs from the suggestion that a situation s is consistent just in case there is no proposition p such that the contradiction $p \ \& \ \neg p$ is true in s:

$$Consistent^{*}(s) \ \equiv_{df} \ \neg\exists p(s \models (p \ \& \ \neg p))$$

[272] Our definition is clearly an object-theoretic counterpart of the classical syntactic definition of a consistent set of sentences Γ. For example, in Enderton 1972 [2001, 134]), we find that a set of sentences Γ is consistent iff there is no formula φ such that both $\Gamma \vdash \varphi$ and $\Gamma \vdash \neg\varphi$.

It is important to recognize that *Consistent* and *Consistent** are independent conditions; there are situations s such that $Consistent(s)$ and $\neg Consistent^*(s)$, and there are situations s' such that $Consistent^*(s')$ and $\neg Consistent(s')$.

To see the former, let q_1 be any proposition and let s_1 be the following canonical situation, which we know exists by (486.3):

$$\imath s\forall p(s \models p \equiv p = (q_1 \,\&\, \neg q_1))$$

Clearly, s_1 is (identical to) a strictly canonical situation, as this was defined at the end of (486); if we let φ be the formula $p = q_1 \,\&\, \neg q_1$, then there is a modally strict proof that $\forall p(\varphi \to \Box\varphi)$. So it follows by modally strict means that exactly one proposition, namely $q_1 \,\&\, \neg q_1$, is true in s_1. Now since every proposition is provably distinct from its negation, it follows that $Consistent(s_1)$, i.e., there is no proposition p such that $s_1 \models p$ and $s_1 \models \neg p$. However, clearly, $\neg Consistent^*(s_1)$, since there is a proposition p, namely q_1, such that $s_1 \models (p \,\&\, \neg p)$.

Now to see that there are situations s' such that $Consistent^*(s')$ and $\neg Consistent(s')$, first note that by (217), we know $\exists p ContingentlyTrue(p)$. Suppose q_2 is such a proposition, so that by definition (213.1), we know both q_2 and $\Diamond\neg q_2$. Now let s_2 be the following situation:

$$\imath s\forall p(s \models p \equiv ((p = q_2) \vee (p = \neg q_2)))$$

s_2 is also (identical to) a strictly canonical situation (exercise). And so we can establish, by modally strict means, that exactly two propositions, namely, q_2 and $\neg q_2$, are true in s_2. But, from what we know about q_2, it is easy to show that there is no proposition p such that $q_2 = (p \,\&\, \neg p)$, and that there is no proposition p such that $\neg q_2 = (p \,\&\, \neg p)$.[273] Hence, it follows that $Consistent^*(s_2)$, since there is no proposition p such that $s_2 \models (p \,\&\, \neg p)$. But $\neg Consistent(s_2)$, since there is a proposition p, namely, q_2, such that both $s_2 \models p$ and $s_2 \models \neg p$.

Thus, *Consistent* and *Consistent** are independent notions. However, as we shall see, the notions are equivalent with respect to actual and possibly actual situations.

(**500**) **Theorem:** Actual Situations are Consistent. If a situation s is actual, then it is consistent:

$$Actual(s) \to Consistent(s)$$

Thus, we've derived what Barwise (1989b, 235) calls the Coherency Principle for actual situations. (We'll introduce a different notion of coherency below.)

Exercise: Using the definition of consistent* in Remark (499), show (a) that actual situations are consistent* and (b) that if s is actual, then s is consistent if and only if s is consistent*.

[273] There can't be a proposition p such that $q_2 = (p \,\&\, \neg p)$ since q_2 is true (by hypothesis) while $p \,\&\, \neg p$ is false for every p (by theorem). Moreover, there can't be a proposition p such that $\neg q_2 = (p \,\&\, \neg p)$ since $\Diamond\neg q_2$ is true (by hypothesis) and $\Diamond(p \,\&\, \neg p)$ is false for every p (by theorem).

(**501**) **Lemmas:** Facts About Consistency. Note also that: (.1) a situation s fails to be consistent if and only if s necessarily fails to be consistent; and (.2) a situation s is possibly consistent if and only if s is consistent:

(.1) $\neg Consistent(s) \equiv \Box\neg Consistent(s)$

(.2) $\Diamond Consistent(s) \equiv Consistent(s)$

(**502**) **Definition:** Possible Situations. We say that a situation s *is possible* iff it is possible that s is actual:

$$Possible(s) \ \equiv_{df} \ \Diamond Actual(s)$$

(**503**) **Theorem:** Facts About Possible Situations. Clearly, (.1) if a situation s is actual, then it is possible; and (.2) if some impossible proposition p is true in s, then s is not possible:

(.1) $Actual(s) \rightarrow Possible(s)$

(.2) $\exists p((s \models p) \,\&\, \neg\Diamond p) \rightarrow \neg Possible(s)$

(**504**) **Theorems:** Consistency and Possible Situations. (.1) Possible situations are consistent; and (.2) there are situations that are consistent but not possible:

(.1) $Possible(s) \rightarrow Consistent(s)$

(.2) $\exists s(Consistent(s) \,\&\, \neg Possible(s))$

One can see why (.2) is a theorem by considering situation s_1 discussed in Remark (499), in which a single proposition, namely $q_1 \,\&\, \neg q_1$, is true. s_1 is consistent, for inconsistency requires that at least two propositions be true in s_1, namely, some proposition and its (distinct) negation. But s_1 is not possible; it couldn't be the case that every proposition true in s_1 is true. A fuller proof is in the Appendix.

Exercise: Using the definition of consistent* in Remark (499), show (a) that possible situations are consistent* and (b) that if s is possible, then s is consistent if and only if s is consistent*.

(**505**) **Exercises:** Sum (Join, Fusion) and Product (Meet) Operations on Situations and a Bounded Lattice. Let us define binary *sum* (*join*, or *fusion*) and *product* (*meet*) operations on situations, $\oplus$ and $\otimes$, respectively, as follows: (.1) $s' \oplus s''$ is the situation s that makes true all and only the propositions p such that either s' or s'' make p true, and (.2) $s' \otimes s''$ is the situation that makes true all and only the propositions p such that both s' and s'' make p true:

(.1) $s' \oplus s'' \ =_{df} \ \imath s \forall p(s \models p \equiv s' \models p \vee s'' \models p))$

(.2) $s' \otimes s'' \ =_{df} \ \imath s\forall p(s \models p \equiv s' \models p \,\&\, s'' \models p))$

So $\oplus$ and $\otimes$ are binary term-forming operation symbols that are well-defined on any pair of *situation* terms. As the first step in this exercise, the reader should confirm that $s' \oplus s''$ and $s' \otimes s''$ are (identical to) strictly canonical situations, i.e., that $\vdash_{\Box} \forall p(\varphi \rightarrow \Box\varphi)$ when (.3) φ is $s' \models p \vee s'' \models p$ and when (.4) φ is $s' \models p \,\&\, s'' \models p$:

(.3) $\vdash_{\Box} \forall p((s' \models p \vee s'' \models p) \rightarrow \Box(s' \models p \vee s'' \models p))$

(.4) $\vdash_{\Box} \forall p((s' \models p \,\&\, s'' \models p) \rightarrow \Box(s' \models p \,\&\, s'' \models p))$

As the next step in the exercise, use (.3) and (.4) to give modally strict proofs of the following laws for join and meet:

(.5) $s \oplus s = s$
$s \otimes s = s$ (Idempotence)

(.6) $s \oplus s' = s' \oplus s$
$s \otimes s' = s' \otimes s$ (Commutativity)

(.7) $s \oplus (s' \oplus s'') = (s \oplus s') \oplus s''$
$s \otimes (s' \otimes s'') = (s \otimes s') \otimes s''$ (Associativity)

(.8) $s \oplus (s \otimes s') = s$
$s \otimes (s \oplus s') = s$ (Absorption)

These laws are characteristic of *lattices*, as defined algebraically. We leave it as a further exercise to show that the situations form a *bounded* lattice, i.e., that $\boldsymbol{s}_{\varnothing}$ and $\boldsymbol{s}_{\boldsymbol{V}}$ serve as the identity elements for $\oplus$ and $\otimes$, respectively, i.e., that:

(.9) $s \oplus \boldsymbol{s}_{\varnothing} = s$

(.10) $s \otimes \boldsymbol{s}_{\boldsymbol{V}} = s$

Finally, the reader may find it interesting to confirm that s is a part of s' if and only if the sum of s and s' just is s':

(.11) $s \trianglelefteq s' \equiv s \oplus s' = s'$

Cf. (629). (.11) is not unrelated to the theorems that show $\trianglelefteq$ to be a reflexive, anti-symmetric and transitive condition on situations (476.1) – (476.3).

We've configured the above as exercises because later we shall develop extend these operations and develop even more general results. The study of (Leibnizian) concepts, in subsection 13.1.4 of Chapter 13, will establish that there are operations $\oplus$ and $\otimes$ that operate on abstract objects generally (though viewed in the guise of concepts). These operations show that the entire domain of abstract objects is structured as a bounded lattice and, indeed, as a Boolean algebra.

(**506**) **Definitions and Theorems:** Notions of Incompatible Situations. Recall that by (199.7), $\overline{p} = \neg p$. With this in mind, we may formulate three different conditions under which situations s and s' are incompatible:

- s and s' are incompatible situations just in case there is a proposition p such that p true in s and $\overline{p}$ is true in s'.
- s and s' are incompatible situations just in case there are propositions p and q such that (a) the conjunction of p and q is impossible, (b) p is true in s, and (c) q is true in s'.
- s and s' are incompatible situations just in case there are propositions p and q such that (a) the conjunction of p and q is false, (b) p is true in s, and (c) q is true in s'.

We may label these notions, respectively, as s and s' are *explicitly*-incompatible (written: s!s'), s and s' are *modally*-incompatible (written: $s \boxtimes s'$), and s and s' are *factually*-incompatible (written: $s|s'$), formalized as follows:

(.1) $s!s' \equiv_{df} \exists p(s \models p \;\&\; s' \models \overline{p})$

(.2) $s \boxtimes s' \equiv_{df} \exists p \exists q(\neg\Diamond(p \;\&\; q) \;\&\; s \models p \;\&\; s' \models q)$

(.3) $s|s' \equiv_{df} \exists p \exists q(\neg(p \;\&\; q) \;\&\; s \models p \;\&\; s' \models q)$

Clearly, (.4) explicit incompatibility implies modal incompatibility, (.5) modal incompatibility implies factual incompatibility, and (.6) explicit incompatibility implies factual incomptability:

(.4) $s!s' \rightarrow s \boxtimes s'$

(.5) $s \boxtimes s' \rightarrow s|s'$

(.6) $s!s' \rightarrow s|s'$

However, (.7) modal incompatibility doesn't imply explicit incompatibility, (.8) factual incompatibility doesn't imply explicit incompatibility, and (.9) factual incompatibility doesn't imply modal incompatibility:

(.7) $\exists s \exists s'(s \boxtimes s' \;\&\; \neg s!s')$

(.8) $\exists s \exists s'(s|s' \;\&\; \neg s!s')$

(.9) $\exists s \exists s'(s|s' \;\&\; \neg s \boxtimes s')$

The proofs in the Appendix construct witnesses to the existential claims.

Exercises: Use definition (505.1) of the sum/join/fusion operation $\oplus$ on situations to show:

$$s!s' \rightarrow (s \oplus s'')!(s' \oplus s''')$$

$$s \boxtimes s' \rightarrow (s \oplus s'') \boxtimes (s' \oplus s''')$$

$$s|s' \rightarrow (s \oplus s'')|(s' \oplus s''')$$

(**507**) **Theorems:** Incompatibility, Consistency, and Possibility. We now examine simple consequences of our definitions that show how the variable forms of incompatibility line up with our notions of consistency and possibility. (.1) if s and s' are explicitly incompatible, then there is no consistent situation of which they are both a part; (.2) if s and s' are modally incompatible, then there is no possible situation in which they are both a part; and (.3) if s and s' are explicitly incompatible, then there is no possible situation of which they are both a part:

(.1) $s!s' \rightarrow \neg\exists s''(Consistent(s'') \,\&\, s \trianglelefteq s' \,\&\, s' \trianglelefteq s'')$

(.2) $s \boxtimes s' \rightarrow \neg\exists s''(Possible(s'') \,\&\, s \trianglelefteq s' \,\&\, s' \trianglelefteq s'')$

(.3) $s!s' \rightarrow \neg\exists s''(Possible(s'') \,\&\, s \trianglelefteq s' \,\&\, s' \trianglelefteq s'')$

12.1.8 The Routley Star Operation on Situations

The Routley 'star' operation was initially introduced in Routley & Routley 1972. Their study of the semantics of entailment assumed the existence of situations ('set-ups') that are neither consistent nor maximal (*ibid*., 335–339).[274]

(**508**) **Definitions and Theorems**. For any situation s, we may define (.1) the Routley star image of s, written s^*, as the situation s' that makes a proposition p true just in case the negation of p fails to be true in s:

(.1) $s^* =_{df} \imath s' \forall p(s' \models p \equiv \neg s \models \overline{p})$

Clearly, for any s, we know $s^*\!\downarrow$, since s^* has been defined in terms of a strictly canonical description. So by familiar, modally strict reasoning it follows that (.2) s^* makes p true just in case $\overline{p}$ fails to be true in s:

[274] Some logicians use the term 'non-normal worlds' to describe situations that are neither maximal (complete) nor consistent. The Routleys, however, used the term 'world' for consistent and maximal situations (1972, 339). In what follows, we reserve the term 'world' for maximal situations, where a maximal situation s is defined in (520) below as one such that, for every proposition p, either s makes p true or s makes the negation of p true. Thus, *possible* worlds will be maximal situations that are possible, while *impossible* worlds will be ones that are not.

(.2) $\forall p(s^* \models p \equiv \neg s \models \overline{p})$

Before we explore the consequences of s^* as just defined, it is important to digress for a moment to consider an alternative definition. One might suppose that instead of s^* being the situation that makes true the propositions whose negations fail to be true in s, one might consider defining s^* as the situation that makes true all the negations of propositions that fail to be true in s, i.e., in terms of the following definition and consequence:

$$s^* =_{df} \imath s' \forall p(s' \models p \equiv \exists q(\neg s \models q \,\&\, p = \overline{q}))$$

$$\forall p(s^* \models p \equiv \exists q(\neg s \models q \,\&\, p = \overline{q}))$$

On this definition, s^* would make p true just in case p is the negation of some proposition that fails to be true in s. Note that in object theory, the condition $\exists q(\neg s \models q \,\&\, p = \overline{q})$ used in this alternative definition is *not* equivalent to the condition $\neg s \models \overline{p}$ used in (.1). The two conditions become equivalent only under the assumption that propositions are identical to their double negations, i.e.,

(ϑ) $\forall p(p = \overline{\overline{p}})$

For note how (ϑ) plays a role in the proof of both directions of the biconditional asserting the equivalence:

$$\exists q(\neg s \models q \,\&\, p = \overline{q}) \equiv \neg(s \models \overline{p})$$

Proof: ($\rightarrow$) Assume $\exists q(\neg s \models q \,\&\, p = \overline{q})$ and suppose it is r, so that $\neg(s \models r) \,\&\, p = \overline{r}$. Now it is easy to show that $\forall p \forall q(p = q \rightarrow \overline{p} = \overline{q})$. So $\neg(s \models r) \,\&\, \overline{p} = \overline{\overline{r}}$. But by ($\vartheta$), it then follows that $\neg(s \models r) \,\&\, \overline{p} = r$. Hence, $\neg(s \models \overline{p})$. ($\leftarrow$) Assume $\neg(s \models \overline{p})$. Then $\neg(s \models \overline{p}) \,\&\, \overline{\overline{p}} = \overline{\overline{p}}$. So, by ($\vartheta$), $\neg(s \models \overline{p}) \,\&\, p = \overline{\overline{p}}$. Existentially generalizing on $\overline{p}$, we obtain $\exists q(\neg(s \models q) \,\&\, p = \overline{q})$.

Since object theory doesn't imply (ϑ) and is consistent with the hyperintensional claim that propositions and their double negations are distinct (albeit necessarily equivalent), we can't use the alternative definition that replaces $\neg s \models \overline{p}$ in (.1) with $\exists q(\neg s \models q \,\&\, p = \overline{q})$.

So to properly capture the definition in Routley & Routley 1972 (338), where they stipulate, that in their canonical model:

> H^* just is the class of propositions A such that $\neg A$ is not in H

we use the condition $\neg s \models \overline{p}$, as we've done in (.1).

We can now easily confirm that (.1) is correct by establishing that s^* behaves as designed. It is straightforward to see that (.3) if s is consistent w.r.t. p (i.e., either p is true in s and $\overline{p}$ is not, or $\overline{p}$ is true in s and p is not), then p is true in s^* if and only if p is true in s:

(.3) $((s \models p \,\&\, \neg s \models \overline{p}) \vee (s \models \overline{p} \,\&\, \neg s \models p)) \rightarrow (s^* \models p \equiv s \models p)$

Next, (.4) a situation is classical w.r.t. double negation (i.e., for every p, s makes p true iff s makes $\overline{\overline{p}}$ true) just in case $s^{**} = s$:

(.4) $\forall p(s \models p \equiv s \models \overline{\overline{p}}) \equiv s^{**} = s$

Cf. Routley & Routley 1972, 338, where it is noted that s is classical ('normal') w.r.t. double negation when $s^{**} = s$.

It also follows that when s is classical w.r.t. double negation, then that s^{**} exhibits additional intended behavior, for (.5) if s is classical w.r.t. double negation, then if s has a 'gap' with respect to q (i.e., $\neg s \models q$ and $\neg s \models \overline{q}$), then s^* has a 'glut' with respect to q (i.e., $s^* \models q$ and $s^* \models \overline{q}$); and (.6) if s is classical w.r.t. double negation, then if s has a glut with respect to q, then s^* has a gap with respect to q:

(.5) $\forall p(s \models p \equiv s \models \overline{\overline{p}}) \rightarrow ((\neg s \models q \,\&\, \neg s \models \overline{q}) \rightarrow (s^* \models q \,\&\, s^* \models \overline{q}))$

(.6) $\forall p(s \models p \equiv s \models \overline{\overline{p}}) \rightarrow ((s \models q \,\&\, s \models \overline{q}) \rightarrow (\neg s^* \models q \,\&\, s^* \models \overline{q}))$

Finally, our definitions imply that (.7) if s is classical w.r.t. double negation, then s is not explicitly incompatible with s^*; (.8) if s is not incompatible with s' and s' is classical w.r.t. double negation, then s' is a part of s^*; and (.9) if s is not incompatible with s' and s' is classical w.r.t. double negation, then the sum/fusion of s' and s^* just is s^*

(.7) $\forall p(s \models p \equiv s \models \overline{\overline{p}}) \rightarrow \neg s!s^*$

(.8) $\neg s!s' \,\&\, \forall p(s' \models p \equiv s' \models \overline{\overline{p}}) \rightarrow s' \trianglelefteq s^*$

Thus, (.7) and (.8) jointly imply that s^* is the 'largest' situation (classical w.r.t. double negation) not explicitly incompatible with s. Finally, for those who completed the Exercises in (505):

Exercise: $\neg s!s' \,\&\, \forall p(s' \models p \equiv s' \models \overline{\overline{p}}) \rightarrow s' \oplus s^* = s^*)$

The reader should confirm that this follows from (.8) by Exercise (505.11).

(**509**) **Remark:** On the Ontology of HYPE. Leitgeb (2019, 321ff) builds a semantics for a system of hyperintensional propositional logic ('HYPE'). He first builds a propositional language $\mathcal{L}$ by starting with atomic propositional letters $p_1, p_2, \ldots$, and logical symbols $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\top$ (where $\rightarrow$ does not express the material conditional). He writes $\overline{p_i}$ for $\neg p_i$, where $\overline{\overline{p_i}}$ is just an abbreviation for p_i. The proposition letters and their negations constitute the *literals*. Leitgeb then constructs HYPE-models for such a propositional language in terms of structures $\langle S, V, \circ, \bot \rangle$, where the elements of the model are simultaneously constrained by the requirements of a Routley star operation *. First, we may describe the elements of the models as follows (Leitgeb 2019, 321–22):

- S is a non-empty set of states.
- V is a function (the valuation function) from S to the power set of the set of literals of the language $\mathcal{L}$, so that each state s in S is associated with a set of literals $V(s)$.
- $\circ$ is a partial fusion function on states that is idempotent and, when defined, commutative and partially associative.
- $\perp$ is a relation of incompatibility that relates states s and s' when some proposition p is true at one and its negation $\overline{p}$ is true at the other.

Moreover, the requirements on the Routley star operation are that:

- $V(s^*) = \{\overline{v} \,|\, v \notin V(s)\}$,
- $s^{**} = s$,
- s and s^* are not incompatible, i.e., $\neg(s \perp s^*)$, and
- s^* is the largest state compatible with s, i.e., if s is not incompatible with s', then the fusion of s' and s^* is defined and the fusion of $s' \circ s^* = s^*$.

In object theory, we may partially reconstruct the elements of HYPE models, in the spirit of footnote 9 (Leitgeb 2019, 323). First we work our way to a definition of a HYPE-state, by first defining atomic-, literal-, and HYPE-propositions:

$$Atomic(p) \equiv_{df} \exists F^n \exists x_1 \ldots \exists x_n (p = [\lambda\, F^n x_1 \ldots x_n])$$

$$Literal(p) \equiv_{df} \exists q(Atomic(q) \,\&\, (p = q \vee p = \overline{q}))$$

$$HYPE(p) \equiv_{df} Literal(p) \,\&\, \overline{\overline{p}} = p$$

$$HYPE\text{-}state(x) \equiv_{df} Situation(x) \,\&\, \forall p(x \models p \rightarrow HYPE(p))$$

So we're identifying HYPE-states not as primitive entities but as situations. Thus when Leitgeb speaks of the members of $V(s)$ as the facts or states of affairs obtaining at s (2019, 322), we may interpret this in terms of our defined notion proposition p is *true in* situation s, as follows:

$$p \in V(s) \equiv_{df} s \models p$$

Note that object theory doesn't guarantee the existence of HYPE-propositions, since it is consistent with the claim that $\overline{\overline{p}} \neq p$. That is, object theory exhibits hyperintensionality at the level of propositions: though $\Box(\overline{\overline{p}} \equiv p)$ is a theorem, it doesn't follow that $\overline{\overline{p}} = p$. So in Leitgeb's initial construction of HYPE models, this kind of hyperintensionality is omitted; later, however, a negated formula

$\neg A$ is defined to be true in a HYPE-state s just in case s is incompatible with every state in which A is true (2019, 326, Df. 5).

Consequently, let us assume, for the remainder of this Remark, that there are HYPE propositions, so as to guarantee the existence of HYPE-states, and let p be a restricted variable over HYPE-propositions and 's' be a restricted variable for HYPE-states (defined above as object-theoretic situations). Then object theory yields comprehension conditions for HYPE-states, which follow *a fortiori* from (486.2):

$$\exists! s \forall p (s \models p \equiv \varphi),$$
provided φ is a condition on HYPE-propositions in which s isn't free

Next, if we put aside, for the moment, the fact that $\circ$ in HYPE is a *partial* binary operation and instead take it to be a total binary fusion operation, then we can model it in terms of the $\oplus$ operation on situations that was defined in (505.1). As Exercises (505.5) – (505.7) demonstrate, $\oplus$ is idempotent, commutative, and associative with respect to the situations. But to model the partiality that $\circ$ exhibits in HYPE, one must introduce a ternary relation R^3 that may or may not relate a pair of situations s and s' to a unique third situation $\iota u Rss'u$. We leave the details for another occasion.[275]

Clearly, we can model the HYPE relation $s \perp s'$ in terms of the *explicit incompatibility* condition $s!s'$. The two principles governing $s \perp s'$ in HYPE (Leitgeb 2019, 322) can be represented in object theory and derived as the following two theorems:

- $(s \models p \,\&\, s' \models \overline{p}) \rightarrow s!s'$
- $s!s' \rightarrow (s \oplus s'')!(s' \circ s''')$

The first follows from the definition of $s!s'$ (506.1), while the second is the first Exercise at the end of (506).

Finally, note that the Routley star operation in defined in HYPE as $V(s^*) = \{\overline{v} \,|\, v \notin V(s)\}$, instead of as $V(s^*) = \{v \,|\, \overline{v} \notin V(s)\}$. However, these two definitions of $V(s^*)$ collapse given that p and $\overline{\overline{p}}$ are identified in the propositional language $\mathcal{L}$ (as noted above). As we saw in the discussion immediately following (508.2), when $\forall p (p = \overline{\overline{p}})$, the definition of s^*, as the situation that makes true all

[275] Intuitively, R would be a partial operation that is idempotent and commutative when $(\iota u Rss'u)\!\downarrow$. Then we could further stipulate that R and the HYPE $\circ$ operation must meet the following conditions:

$$s \circ s' =_{df} \iota s'' \forall p (s'' \models p \;\equiv\; s \models p \vee s' \models p \vee \iota u (Rss'u) \models p)$$

$$\iota s_4 R s_1 ((s_1 \circ s_2) \circ s_3) s_4 \trianglelefteq (s_1 \circ s_2) \circ s_3$$

The intuition here is that R ensures that $s \circ s'$ contains propositions other than the ones that are in s and s'. Constraints on R make the conditions (idempotence, commutativity when defined, partial associativity when defined) true. The extra constraint on R guarantees partial associativity.

the propositions whose negations fail to be true in s, becomes equivalent to the definition of s^* as the situation that makes true all of the negations of propositions that fail to be true in s. So, by identifying $p = \overline{\overline{p}}$ in the HYPE language $\mathcal{L}$, HYPE models exhibit $\forall p(p \in V(s) \equiv \overline{\overline{p}} \in V(s))$. In object-theoretic terms, this condition becomes $\forall p(s \models p \equiv s \models \overline{\overline{p}})$. As we saw in (508.4), this last condition is equivalent to $s^{**} = s$.

Thus, the Routley star operation in HYPE can by modeled by the operation s^* defined in (508.1). Under the stipulation that $s^{**} = s$, the theorems governing s^* in (508.3) – (508.7) match the requirements of the Routley star operation defined in Leitgeb 2019 (322), modulo the partiality of the HYPE $\circ$ operation.

12.2 Possible Worlds

(**510**) **Remark:** On the Nature of Possible Worlds. It is only relatively recently that philosophers and logicians have started to think seriously and systematically about the nature of possible worlds. Of course, Leibniz mentioned them in both his *Theodicy* (T 128 = G.vi 107) and in the *Monadology* §53 (PW 187 = G.vi 615–616). But contemporary philosophers have expressed skepticism about Leibniz's conception of possible worlds, as in the following passage in Stalnaker 1976 (65):[276]

> According to Leibniz, the universe—the actual world—is one of an infinite number of possible worlds existing in the mind of God. God created the universe by actualizing one of these possible worlds—the best one. It is a striking image, this picture of an infinite swarm of total universes, each by its natural inclination for existence striving for a position that can be occupied by only one, with God, in his infinite wisdom and benevolence, settling the competition by selecting the most worthy candidate. But in these enlightened times, we find it difficult to take this metaphysical myth any more seriously than the other less abstract creation stories told by our primitive ancestors. Even the more recent expurgated versions of the story, leaving out God and the notoriously chauvinistic thesis that our world is better than all the rest, are generally regarded, at best, as fanciful metaphors for a more sober reality.

So let's, for the moment, suspend judgment about Leibniz's conception of possible worlds, and instead focus next on Wittgenstein's famous characterization of the actual world in the opening lines of his *Tractatus* (1921), where we find the claims "The world is all that is the case" and "The world is the totality of

[276]One shouldn't conclude on the basis of this passage that Stalnaker is a possible worlds skeptic, since he goes on to employ possible worlds in his work on the analysis of counterfactual conditionals. He just takes them to be properties of a special kind. See the discussion below.

facts, not of things" (1921, 7, Propositions 1 and 1.1). It is unclear whether Wittgenstein had a view about nonactual possible worlds; subsequent propositions in the *Tractatus* mention *possible states of affairs* but the text doesn't explicitly say whether any are so total and complete as to constitute a possible world (see, e.g., Propositions 2.012, 2.0124, and 2.013 in the *Tractatus*). Interestingly, Carnap (1947, 9) interpreted Wittgenstein's text as if some are.

In what follows, we shall be attempting to prove theorems about the *nature* of possible worlds; we shall not be interested in *models* of them. We take it that possible worlds are *none of the following*: (a) *complete and consistent sets of sentences*, of the kind described in Lindenbaum's Lemma (Tarski 1930, 34, fn. ‡), (b) *state descriptions*, as defined by Carnap (1947, 9), (c) *model sets*, as defined by Hintikka (1955), (d) *variable assignments* agreeing on the values of the individual variables, as defined by Kripke (1959, 2–3), or (e) *Tarski models*, as put to use in Montague 1960. Similarly, I take the following works to contain mathematical models, not theories, of possible worlds: (f) Pollock 1967 (317), in which a possible world is identified as any *set* of states of affairs that is maximal and consistent, (g) Quine 1968 (14–16), in which possible worlds are identified with *sets* of quadruples of real numbers representing the coordinates of spacetime points occupied by matter, (h) Cresswell 1972a (6), in which possible worlds are identified as *sets* of basic particular situations, (i) Adams 1974 (225), in which talk of possible worlds is reduced to talk about maximal, consistent *sets* of propositions ('world-stories'), and (j) Menzel 1990 (371ff), in which appropriately structured Tarski models represent possible worlds in virtue of having the modal property *possibly being a representation of the way things are*. The proposals mentioned in (a) – (j) fail to be theories of possible worlds because possible worlds, whatever they are, are not mathematical objects.[277] They are not sets of formal sentences, sets of propositions, sets of states of affairs, formal models, model sets, or assignments to variables. Of course, for some purposes, these mathematical objects might serve to *represent* possible worlds, but our interest is in the worlds as objects in their own right.

Although the notion of *possible world* gained currency in the 1950s, the philosophers and logicians in that decade weren't especially interested in their

[277] The physicist Mark Tegmark (2008) asserts that the physical universe is a mathematical structure. Of course, this modern-day Pythagoreanism requires an analysis of what a mathematical structure is, something Tegmark doesn't provide. By contrast, we shall analyze mathematical structures as certain abstract objects. See Chapter 15. Hence, even if Tegmark is right about what the physical universe is, the present theory offers a further analysis of the mathematical structure in question.

For the present purposes, however, it suffices to note that Tegmark's conception of a physical universe as a mathematical structure doesn't lead to a theory of worlds as mathematical structures. Possible worlds are objects whose identity is, in part, tied to the non-mathematical propositions that are true or false at them; mathematical structures are just not the kind of thing at which non-mathematical propositions can be true or false.

nature as entities in their own right. Copeland (2006, 381) describes a letter from Carew Meredith to Arthur Prior, dated 10 October 1956, in which Meredith uses the term 'possible world' when demonstrating how to falsify a certain formula. Copeland also describes a 1957 lecture handout from Timothy Smiley at the University of Cambridge, which indicates that necessary truths are true in all 'possible worlds', and a proposition is possible if it is true in *some* 'possible world' (2002, 121; 2006, 385; quotes in the original).[278] Moreover, Hintikka (1957, 61–62) suggested that models and model sets:

> ...correspond to the different situations we want to consider in modal logic, and they are interconnected, in the first place, by a rule saying (roughly) that whatever is necessarily true in the actual state of affairs must be (simply) true in all the alternative states of affairs.

The text doesn't make it clear whether "alternative states of affairs" can be partial (i.e., similar to situations) or are always total (and hence alternative possible worlds). Finally, though Kripke (1959) interprets modal logic in terms of a set of alternative assignments to variables, he then says (1959, 2):

> The basis of the informal analysis which motivated these definitions is that a proposition is necessary if and only if it is true in all "possible worlds". (It is not necessary for our present purposes to analyze the notion of a "possible world" any further.)

These invocations of possible worlds played an important role in the development of the semantics of modal logic in the early 1960s (Kripke 1963a, 1963b; Prior 1963). In all of these cases, however, it is fair to say that modal logicians weren't yet interested in the *nature* of the possible worlds, even though they thought that some such notion helps us to understand formal frameworks for interpreting modal language.

The study of the nature of possible worlds began in earnest when Lewis (1968) introduced *axioms* governing worlds. Though Lewis was, in the 1968 paper, primarily interested in formulating his 'counterpart theory' of possibilia, his axioms used variables ranging over worlds and implicitly defined properties he took worlds to have. His work gave rise to the goal of precisely specifying the nature of possible worlds, however, and theories of them were subsequently developed in a variety of other works, including Lewis 1973; Plantinga 1974, 1976; Fine 1977; Chisholm 1981; Zalta 1983; Pollock 1984; Lewis 1986; Armstrong 1989, 1997; and Stalnaker 2012.[279]

[278] In Copeland 2002, footnote 25, and again in Copeland 2006, footnote 18, we find an acknowledgment to David Shoesmith for providing a copy of the lecture handout. I've not been able to acquire a copy.

[279] Interestingly, Stalnaker's position, until 2012, was that one can assume the existence of pos-

It is worthwhile recalling a famous passage in the work of David Lewis that contains a kind of *credo* and justification for the belief in possible worlds (Lewis 1973, 84):

> I believe that there are possible worlds other than the one we happen to inhabit. If an argument is wanted, it is this. It is uncontroversially true that things might be otherwise than they are. I believe, and so do you, that things could have been different in countless ways. But what does this mean? Ordinary language permits the paraphrase: there are many ways things could have been besides the way they actually are. On the face of it, the sentence is an existential quantification. It says that there exist many entities of a certain description, to whit 'ways things could have been'. I believe things could have been in countless ways; I believe permissible paraphrases of what I believe; taking the paraphrase at its face value, I therefore believe in the existence of entities that might be called 'ways things could have been'. I prefer to call them possible worlds.

Lewis, in a later work, offers a further justification for belief in possible worlds. He writes (1986, 3):

> Why believe in a plurality of worlds? – Because the hypothesis is serviceable, and that is a reason to think that it is true. The familiar analysis of necessity as truth at all possible worlds was only the beginning. In the last two decades, philosophers have offered a great many more analyses that make reference to possible worlds, or to possible individuals that inhabit

---

sible worlds in semantic analysis without endorsing any particular theory about their nature. In 1984 (Chapter 3), the final paragraph of his 1976 paper is revised and expanded, to include the following (57):

> ...the moderate realism I want to defend need not take possible worlds to be among the ultimate furniture of the world. Possible worlds are primitive notions of the theory, not because of their ontological status, but because it is useful to theorize at a certain level of abstraction, a level that brings out what is common in a certain range of otherwise diverse activities. The concept of possible worlds that I am defending is not a metaphysical conception, although one application of the notion is to provide a framework for metaphysical theorizing. The concept is a formal or functional notion, like the notion of *individual* presupposed by the semantics for extensional quantification theory. ...
>
> Similarly, a possible world is not a particular kind of thing or place. The theory leaves the *nature* of possible worlds as open as extensional semantics leaves the nature of individuals. A possible world is what truth is relative to, what people distinguish between in their rational activities. To believe in possible worlds is to believe only that those activities have a certain structure, the structure which possible worlds theory helps to bring out.

But in 2012 (8), Stalnaker puts forward the suggestion that possible worlds are properties, i.e., *ways a world might be*. So a possible world is "the kind of thing that is, or can be, instantiated or exemplified" (2012, 8).

> possible worlds. I find that record most impressive. I think it is clear that talk of *possibilia* has clarified questions in many parts of the philosophy of logic, mind, of language, and of science – not to mention metaphysics itself. Even those who officially scoff often cannot resist the temptation to help themselves abashedly to this useful way of speaking.

And on the page before, Lewis states what seems to be the most important principle governing worlds, namely, "... absolutely *every* way that a world could possibly be is a way that some world *is*" (1986, 2, [emphasis in the original]).

It is not surprising that Lewis' suggestion, that possible worlds are physically-disconnected concrete entities inhabiting some logical space, was controversial. Van Inwagen 1986 (185–6) contrasts two of the leading conceptions of worlds that emerged:

> Lewis did not content himself with saying that there were entities properly called 'ways things could have been'; ... He went on to say that what most of us would call 'the universe', the mereological sum of all the furniture of earth and the choir of heaven, is one among others of these 'possible worlds' or 'ways things could have been', and that the others differ from it "not in kind but only in what goes on in them" (Lewis 1973, 85).
>
> Whether or not the existence of a plurality of universes can be so easily established, the thesis that possible worlds are universes is one of the two 'concepts of possible worlds' that I mean to discuss. ... The other concept I shall discuss is that employed by various philosophers who would probably regard themselves as constituting the Sensible Party: Saul Kripke, Robert Stalnaker, Robert Adams, R.M. Chisholm, John Pollock, and Alvin Plantinga.[3] These philosophers regard possible worlds as abstract objects of some sort: possible histories of *the* world, for example, or perhaps properties, propositions or states of affairs.

Van Inwagen goes on to call Lewis a 'concretist' about worlds while the members of the 'Sensible Party' are called 'abstractionists'. The main difference between these two conceptions is whether worlds are to be defined primarily in terms of a part-whole relation (concretists) or in terms of the propositions true at them (abstractionists).

Given this opposition, it seems that the abstractionist conception is a kind of generalization of Wittgenstein's view of the actual world in the *Tractatus*, mentioned above. But Menzel (2015) notes that there is a third important conception of worlds, namely, the 'combinatorialist' conception, on which possible worlds are taken to be "recombinations, or rearrangements, of certain metaphysical simples," where "both the nature of simples and the nature of recombination vary from theory to theory" (Section 2.3). This view is exhibited in the work of Quine, Cresswell, and Armstrong, *op. cit.*. On Menzel's analysis, Wittgenstein's view is more closely allied with the combinatorialist conception than the abstractionist one.[280] Of course, if one defines *actualism* in some other way (i.e., other than by eschewing possible but non-actual objects), then the present theory might fail to be actualist.

With these introductory remarks, we may turn to the subtheory of possible worlds that is developed in object theory. I take the theory described below to be unique in that the principles governing worlds are precisely derived as theorems rather than stipulated as axioms. Some of the theorems below may already be familiar, having appeared in Zalta 1983 (78–84), 1993, Fitelson & Zalta 2007, Bueno, Menzel & Zalta 2014, and Menzel & Zalta 2014. Many of these theorems have been reworked and enhanced in several ways. Moreover, the theorems now have greater significance, since object theory now asserts that *every* formula signifies a proposition (104.2), including those with encoding subformulas or encoding formulas as subterms. So when the theory implies, for an arbitrary possible world w, that w is maximal (i.e., implies that for every proposition p, either p is true at w or $\neg p$ is true at w), every formula can be instantiated for the universal quantifier over all propositions.[281]

Even with these improvements, the basic idea has remained the same: possible worlds are defined to be situations of a certain kind and, hence, abstract individuals.[282] Thus, our conception of worlds can be traced back to Wittgen-

---

[280] See again Menzel 1990, where a way that philosophers might do without possible worlds altogether is suggested. Menzel develops a homophonic interpretation of modal language that may suffice for the purposes of analyzing our modal beliefs. Menzel suggests that we use certain Tarski models to represent the possible worlds of Kripke models and then he attributes modal properties to them. On Menzel's proposal, a Tarski model meeting certain conditions *could have represented* the world as it would have been had things been different. One could put the view more simply as follows: each such Tarski model might have been a model of the actual world. So not only does Menzel avoid claiming that Tarski models are possible worlds, strictly speaking he also avoids the claim that the former represent the latter, since on his view, we can simply rest with the primitive modal properties of Tarski models and thereby provide a precise, but homophonic, interpretation of modality.

Linsky and Zalta (1994, 444) raised some concerns about this suggestion. I only add here that if Menzel's goal of doing without possible worlds is to avoid the committment to 'possible but non-actual objects', then the theory of possible worlds developed below achieves this goal, given the Quinean interpretation of the quantifiers of our formalism. For on that interpretation, our possible worlds are existing, actual (abstract) objects and are 'possible' only in a defined sense that is consistent with actualism; see theorem (517) below and the definitions in terms of which it is cast.

[281] When we move to typed object theory in Chapter 15, there will be some formulas that denote nothing. For example, the definite description $\imath p(p \,\&\, \neg p)$, where p is a proposition variable, becomes expressible. This description is both a term and a formula. As a term it denotes nothing, but as a formula, it has truth conditions. So, in what follows, when we remark upon the fact that a theorem of world theory expressed with propositional variables holds for arbitrary formulas, this only applies to second-order object theory. In typed object theory, we can instantiate such theorems only to significant formulas.

[282] Our theory therefore reconciles situation theory and world theory. In the early 1980s, situation theory was thought to be incompatible with world theory. See the early publications of Barwise

stein, but with the added insight about the distinction between the internal and external properties of situations discussed in Remark (466). The Tractarian conception "the world is all that is the case" will be validated by the theorem that p is true if and only if p is true at the actual world (536.2)★, since this theorem implies that the actual world *encodes* all that is the case. The Tractarian conception will be preserved in our definition of a possible world (512) as a situation s that might be such that all and only true propositions are true in s, since this definition implies that a possible world is a situation that might be such that it *encodes* everything that is the case. But we begin our study of possible worlds with some important lemmas about a key group of situations.

(**511**) **Lemmas:** Situations In Which All and Only True Propositions are True. We may import and export the classical connectives on the basis of the following theorems. (.1) If all and only true propositions are true in s, then for every proposition q, $\neg q$ is true in s if and only if it is not the case that q is true in s; and (.2) if all and only true propositions are true in s, then for every proposition q and r, $q \rightarrow r$ is true in s if and only if, if q is true in s then r is true in s:

(.1) $\forall p(s \models p \equiv p) \rightarrow \forall q((s \models \neg q) \equiv \neg(s \models q))$

(.2) $\forall p(s \models p \equiv p) \rightarrow \forall q \forall r((s \models (q \rightarrow r)) \equiv ((s \models q) \rightarrow (s \models r)))$

Now if we switch to theorem schemata, so as not to have to worry about captured variables, we have (.3) if all and only true propositions are true in s, then every-α-is-such-that-φ is true in s if and only if, every α is such that φ-is-true-in-s:[283]

(.3) $\forall p(s \models p \equiv p) \rightarrow ((s \models \forall \alpha \varphi) \equiv \forall \alpha(s \models \varphi))$

Clearly if we let φ be the variable q, then (.3), GEN, and (39.3) imply that if all and only true propositions are true in s, then for every proposition q, $\forall \alpha q$ is true in s iff for every α, q is true in s.

Turning now to modality, we switch back to variables ranging over propositions. We now show that if s makes true all and only true propositions, then the necessity operator can only be exported from, not imported into, a *truth-in-s* context. For the existence of contingently true propositions undermines the

and Perry, *op. cit.*

[283]If we had formulated (.3) as:

$$\forall p(s \models p \equiv p) \rightarrow ((s \models \forall \alpha q) \equiv \forall \alpha(s \models q))$$

and then universally generalized on q, then even though $\varphi\downarrow$ holds for every formula φ, the only formulas we could substitute for $\forall q$ are those in which α doesn't occur free. So if a formula φ has α free, we'd have to switch to an alphabetic variant of the above, e.g.,

$$\forall p(s \models p \equiv p) \rightarrow ((s \models \forall \beta q) \equiv \forall \beta(s \models q))$$

and universally generalize on q in order to instantiate to φ. But by formulating (.3) as in the text, the theorem holds straightaway for any formula.

importation. We thus have following theorems: (.4) if all and only true propositions are true in s, then if necessarily-q is true in s, then necessarily q-is-true-in-s; and (.5) if all and only true propositions are true in s, then for some proposition q, then necessarily, q is true in s while it is not the case that $\Box q$ is true in s:

(.4) $\forall p(s \models p \equiv p) \rightarrow \forall q((s \models \Box q) \rightarrow \Box(s \models q))$

(.5) $\forall p(s \models p \equiv p) \rightarrow \exists q(\Box(s \models q) \,\&\, \neg(s \models \Box q))$

Exercise: Determine whether the following is provable and, if so, whether there is a modally strict proof: $\forall p(s \models p \equiv p) \rightarrow \forall q((s \models \mathcal{A}q) \equiv \mathcal{A}(s \models q))$.

Finally, we establish that (.6) there is a situation that makes true all and only true propositions:

(.6) $\exists s \forall p(s \models p \equiv p)$

Situations that *might* make true all and only truths will now be identified as *possible worlds*.

(**512**) **Definition and Theorems:** Possible Worlds. Recall that in definition (470), we stipulated that a proposition p is true in x whenever x is a situation and x encodes p, where the latter was defined as $x[\lambda y\, p]$ (295). Then since situations are abstract objects, we now say that an object x is a *possible world* iff x is a situation that is possibly such that all and only the propositions true in x are true:

(.1) $PossibleWorld(x) \equiv_{df} Situation(x) \,\&\, \Diamond\forall p(x \models p \equiv p)$

Note that our definition decides Choice 3 in Barwise 1989a (261) in favor of Alternative 3.1: worlds are situations. The question may arise, why haven't we defined possible worlds as maximal and consistent situations? We take up this question in (523), after we formulate the definition of *maximality* and prove, in (521) and (518) below, that possible worlds are maximal and consistent.

To see that $PossibleWorld(x)$ is a *restriction condition*, as this metatheoretic notion was defined in (336), we first establish, as a modally strict theorem, that (.2) there exist possible worlds, and then establish, as a modally strict theorem schema, (.3) if κ is a possible world, then κ is significant, for any individual term κ:

(.2) $\exists x PossibleWorld(x)$

(.3) $PossibleWorld(\kappa) \rightarrow \kappa\!\downarrow$

(.2) tells us that $PossibleWorld(x)$ is strictly non-empty, and (.3) tells us that $PossibleWorld(x)$ has strict existential import. Thus, by (336), $PossibleWorld(x)$ is a restriction condition. But before we introduce restricted variables, we establish facts that show $PossibleWorld(x)$ is a rigid restriction condition.

(**513**) **Theorem:** Rigidity of *PossibleWorld*(x). Our definitions imply that (.1) x is a possible world if and only if necessarily x is a possible world; (.2) possibly x is a possible world if and only if x is a possible world; and (.3) possibly x is a possible world if and only if necessarily x is a possible world; and (.4) actually x is a possible world if and only if x is a possible world:

(.1) $PossibleWorld(x) \equiv \Box PossibleWorld(x)$

(.2) $\Diamond PossibleWorld(x) \equiv PossibleWorld(x)$

(.3) $\Diamond PossibleWorld(x) \equiv \Box PossibleWorld(x)$

(.4) $\mathcal{A}PossibleWorld(x) \equiv PossibleWorld(x)$

(**514**) **Remark:** Restricted Variables for Possible Worlds. If we apply GEN to the left-to-right direction of (513.1), then we know:

$$\vdash_{\Box} \forall x(PossibleWorld(x) \rightarrow \Box PossibleWorld(x))$$

Thus, *PossibleWorld*(x) is not only a restriction condition in the sense of (336) but a rigid restriction condition in the sense of (340). So let $w, w', w'', \ldots$ be rigid, restricted variables that range over any x such that *PossibleWorld*(x). We may use these variables according to the conventions described in (337) – (341).

Note that our restricted variables for possible worlds present us with two interpretive options: we may regard them either as singly-restricted variables ranging over objects x such that *PossibleWorld*(x) or we may regard them as doubly restricted variables that range over the situations s such that *Possible-World*(s). The latter option has some advantages: by using the rigid restricted variable s for situations, the definition of possible world in (512) can be more simply expressed as follows, given our conventions for free restricted variables in definitions:

$$PossibleWorld(s) \equiv_{df} \Diamond \forall p(s \models p \equiv p)$$

Moreover, by regarding $w, w', \ldots$ as doubly restricted variables that range over situations, we can immediately regard notions defined with respect to situations (using a free restricted variable s) as defined with respect to possible worlds w. For example, since *Consistent*(s) is defined with respect to situations (498), the claim *Consistent*(w) doesn't need to be defined. See the next Remark (515) for another example.

Of course, we are free to exercise either interpretation of the rigid restricted variables $w, w', \ldots$ when it is convenient to do so. But note that we now have *two* ways of eliminating the bound restricted variable w. Consider just the variable-binding sentence-forming operators. Then if w is interpreted as singly restricted, then we may expand the claims:

(A) $\forall w Consistent(w)$

(B) $\exists w Actual(w)$

as the claims:

(C) $\forall x(PossibleWorld(x) \to Consistent(x))$

(D) $\exists x(PossibleWorld(x) \,\&\, Actual(x))$

But if w is interpreted as doubly restricted, i.e., as a restricted variable ranging over situations s such that $PossibleWorld(s)$, then (A) and (B) may be interpreted as:

(E) $\forall s(PossibleWorld(s) \to Consistent(s))$

(F) $\exists s(PossibleWorld(s) \,\&\, Actual(s))$

Of course (E) and (F) have bound restricted variables which in turn may be eliminated.[284]

Another advantage of regarding w as doubly restricted is that by treating (A) and (B) as shorthand, respectively, for (E) and (F) instead of as shorthand for (C) and (D), we can often simplify the reasoning process: the consequences of (E) and (F) obtained by Rules ∀E and ∃E, for example, will involve terms that are known or assumed to refer rigidly to situations.

Consequently, in what follows, we sometimes prefer to reason as if our restricted variables for possible worlds are doubly restricted. Of course, in some cases, we continue to reason by eliminating the possible world variables in the usual way, as singly restricted. We do this on those occasions when it clarifies the reasoning using a principle (axiom, theorem, or definition) for which it may not be immediately evident how it applies to restricted variables.

Moreover, in what follows, the reader may encounter a series of principles (i.e., definitions and theorems) in which (a) the definition introduces some new notion defined on situations using the variable s while (b) the subsequent theorem governs possible worlds and is stated using the variable w. The reason for this should now be clear: in general, we always choose the variable, restricted or otherwise, that is the most appropriate for the occasion, i.e., the variable that either (i) yields the simplest formulation of the definition and the simplest statement of the theorem (without compromising generality), or (ii) minimizes the amount of reasoning that has to be done to prove the theorem (thereby

[284] In the usual way, (E) and (F) expand to:

$$\forall x(Situation(x) \to (PossibleWorld(x) \to Consistent(x)))$$

$$\exists x(Situation(x) \,\&\, PossibleWorld(x) \,\&\, Actual(x))$$

Note, however, that we shall not have occasion to regard (A) and (B), respectively, as short for the above, since the clause $Situation(x)$ is redundant in both.

minimizing the chances of introducing reasoning errors and maximizing the clarity of the justification).

(**515**) **Remark:** Truth At a Possible World. In the previous Remark, we noted that if possible world variables are interpreted as doubly restricted, i.e., as ranging over situations s such that $PossibleWorld(s)$, then notions defined on situations apply to possible worlds without having to be redefined. An important example is the notion of *truth in* a situation, i.e., $s \models p$. By interpreting w as a doubly restricted variable, one can see that $w \models p$ is defined. Note, however, that whereas we read $s \models p$ as 'p is *true in s*', we shall, for historical reasons, read $w \models p$ as 'p is *true at w*'. Thus, *truth at* a possible world is simply a special case of the notion *truth in* a situation. Moreover, since worlds are a special type of situation, it follows by theorem (471) that $w \models p$ is equivalent to $w[\lambda y\, p]$. Hence, when a proposition p is true at w, the property $[\lambda y\, p]$ *characterizes* w by way of an encoding predication.

(**516**) **Theorem:** Identity Conditions for Possible Worlds. It is a consequence of the fact that possible worlds are situations that w and w' are identical just in case they make the same propositions true:

$$w = w' \equiv \forall p(w \models p \equiv w' \models p)$$

This theorem validates a pre-theoretic intuition about Wittgensteinian possible worlds and it does so (a) by proof rather than by stipulation and (b) without identifying possible worlds as sets. Though possible worlds have extensional identity conditions, they are nevertheless hyper-hyperintensional objects; they are even more fine-grained than the hyperintensional propositions they encode.

(**517**) **Theorem:** A Non-Trivial Theorem. It is now straightforward to show that possible worlds are possible:

$$Possible(w)$$

The reader is encouraged to think twice about this theorem. It is not a triviality; the claim to be established is not a tautology of the form $(\varphi \,\&\, \psi) \rightarrow \varphi$. We're not proving $Possible(s) \,\&\, World(s) \rightarrow Possible(s)$. A second look reveals that, given the definition of $Possible(s)$ (502) as $\Diamond Actual(s)$, the theorem captures a fundamental intuition about possible worlds, namely, that each world *w might have been actual*, i.e., might have been such that every proposition true at w is true. This claim is so fundamental to our conception of possible worlds that we tend to overlook the suggestion that it is capable of proof. And, just as importantly, we've *not* used a primitive notion of possibility to assert a modal claim about a primitive notion of possible world. It would be a mistake to use the primitive notions of *possibility* and *possible world* to simply assert that possible worlds might have been actual. Such a move borders on nonsense: since the primitive, *non-modal* concept of a possible world was (historically) introduced to interpret primitive modal claims, the use of a primitive modal notion to assert modal claims about primitive possible worlds seems confused at best, and nonsense at worst. By contrast, we are using a single modal primitive ($\Box$) and its defined dual ($\Diamond$), and the *defined* conditions $PossibleWorld(x)$ and $Possible(x)$, to derive an fact that many philosophers assume when reasoning about possible worlds. So, this theorem has greater significance than meets the eye.

(**518**) **Theorem:** Possible Worlds Are Consistent and Non-Trivial. Clearly, (.1) possible worlds are consistent, and (.2) non-trivial situations:

(.1) $Consistent(w)$

(.2) $\neg TrivialSituation(w)$

Exercises: (a) Prove that $Consistent^*(w)$, as this was defined in Remark (499). (b) Prove that possible worlds are not null situations, as the latter were defined in (487.1).

(**519**) **Lemmas:** Rigidity of Truth At a Possible World. It is straightforward to establish that (.1) p is true at w if and only if necessarily p-is-true-at-w; (.2) possibly p-is-true-at-w if and only if p is true at w; (.3) possibly p-is-true-at-w if and only if necessarily p-is-true-at-w; (.4) actually p-is-true-at-w if and only if p is true at w; and (.5) p fails to be true-at-w if and only if necessarily p fails to be true-at-w:

(.1) $w \models p \equiv \Box w \models p$

(.2) $\Diamond w \models p \equiv w \models p$

(.3) $\Diamond w \models p \equiv \Box w \models p$

(.4) $\mathcal{A} w \models p \equiv w \models p$

(.5) $\neg w \models p \equiv \Box \neg w \models p$

(**520**) **Definition:** Maximality. A situation s is *maximal* iff for every proposition p, either p is true in s or $\neg p$ is true in s:

$$Maximal(s) \equiv_{df} \forall p(s \models p \lor s \models \neg p)$$

(**521**) **Theorem:** Possible Worlds Are Maximal.

$$Maximal(w)$$

We've therefore now established that possible worlds are maximal and consistent.

(**522**) **Theorem:** Maximality, Possible Worlds, and Possibility.

(.1) $Maximal(s) \rightarrow \Box Maximal(s)$

(.2) $PossibleWorld(s) \equiv Maximal(s) \,\&\, Possible(s)$

(**523**) **Remark:** Possible Worlds as MaxCon Situations. It would serve well to remark upon the difference between defining possible worlds as in (512) and defining them as *maxcon* situations, i.e., as situations that are maximal, consistent, and consistent*. The simple answer is that the definitions aren't equivalent; whereas situations satisfying (512) are maxcon, it doesn't follow that situations that are maxcon satisfy (512). To see this, consider a maxcon situation in which (it is true that) Socrates might have been a chunk of rock or might have had different parents. If certain essentialist principles are true and it is a metaphysical fact that Socrates couldn't have been a chunk of rock and couldn't have had different parents, then a maxcon situation in which Socrates is a chunk of rock or had different parents is one that fails to be a possible world in the sense of (512). In other words, there may be modal facts that place constraints on what maxcon situations are possible worlds. (512) defines *possible world* in a way that respects those modal truths, should there be any.

Moreover, theorem (522.2) shows that definition (512) implies that a situation is a possible world if and only if it is maximal and might have been actual, where the modality 'might' is metaphysical possibility. This is not captured by saying that that possible worlds are situations that are maximal and free of contradictions. Finally, the body of theorems derivable from (512) and its supporting definitions provide evidence of its power and correctness. This evidence will continue to accumulate as we go.

(**524**) **Definition:** Necessary Implication and Necessary Equivalence. We say that p *necessarily implies* q, written $p \Rightarrow q$, just in case necessarily, if p then q:

(.1) $p \Rightarrow q \;\equiv_{df}\; \Box(p \rightarrow q)$

We also say that p and q are *necessarily equivalent*, written $p \Leftrightarrow q$, just in case both p necessarily implies q and q necessarily implies p:

(.2) $p \Leftrightarrow q \;\equiv_{df}\; (p \Rightarrow q) \,\&\, (q \Rightarrow p)$

Clearly, then, we have extended the notions of *necessary implication* and *necessary equivalence*, defined in (442.1) and (442.2), respectively, so that they apply to propositions as well as properties.

(**525**) **Theorem:** Necessary Equivalence and Necessary Material Equivalence. It is therefore a simple consequence of the previous definition that p and q are necessarily equivalent if and only if necessarily, p is materially equivalent to q:

$$p \Leftrightarrow q \equiv \Box(p \equiv q)$$

(**526**) **Lemmas:** Facts About Situations. Let us now use $p_1, \ldots, p_n$ (which we've heretofore taken to be constants) as *variables* ranging over propositions. Then we may establish two helpful lemmas about situations. Where $n \geq 1$, a theorem of the following form holds, namely, (.1) if $p_1, \ldots, p_n$ are all true in situation s and the conjunction of $p_1, \ldots, p_n$ implies q, then if all and only true propositions are true in s, then q is true in s:

(.1) $(s \models p_1 \ \& \ \ldots \ \& \ s \models p_n \ \& \ ((p_1 \ \& \ \ldots \ \& \ p_n) \rightarrow q)) \rightarrow (\forall p(s \models p \equiv p) \rightarrow s \models q)$ $(n \geq 1)$

Since these are modally strict theorems, the Rule of Necessitation (RN) applies to each. Hence the necessitation of each n-ary instance of (.1) is a theorem, $n \geq 1$ So by the K axiom and other principles of modality, we can establish theorems of the form (.2) if necessarily p_1-is-true-in-s and … and necessarily p_n-is-true-in-s and $p_1 \& \ldots \& p_n$ necessarily implies q, then necessarily, q-is-true-in-s when all and only truths are true in s:

(.2) $(\Box s \models p_1 \ \& \ \ldots \ \& \ \Box s \models p_n \ \& \ ((p_1 \ \& \ \ldots \ \& \ p_n) \Rightarrow q)) \rightarrow$
$\Box(\forall p(s \models p \equiv p) \rightarrow s \models q)$ $(n \geq 1)$

(**527**) **Definition:** n-Modal Closure, i.e., n-ary Closure Under Necessary Implication. We now say, for any choice of $n \geq 1$, that a situation s is *n-modally closed* or *n-ary-closed under necessary implication* if and only if, for any propositions $p_1, \ldots, p_n$ and q, if each of $p_1, \ldots, p_n$ is true in s and q is necessarily implied by the conjunction $p_1 \ \& \ \ldots \ \& \ p_n$, then q is also true in s:

$$\begin{array}{l} n\text{-}ModallyClosed(s) \equiv_{df} \\ \quad \forall p_1 \ldots \forall p_n \forall q((s \models p_1 \ \& \ \ldots \ \& \ s \models p_n \ \& \ ((p_1 \ \& \ \ldots \ \& \ p_n) \Rightarrow q)) \rightarrow s \models q) \end{array}$$

Though one might suggest that it is the notion of *truth in a situation* that is n-modally closed when the definiens obtains, this definition does no harm: it is reasonable to speak as if situations are n-modally closed if the above condition holds.

(**528**) **Theorem:** Possible Worlds are n-Modally Closed. From (526) and (527), it follows that possible worlds are n-modally closed, for any choice of n:

$$n\text{-}ModallyClosed(w) \qquad (n \geq 1)$$

(**529**) **Corollary:** Negation Coherence of Truth At A World. We may now establish that truth-at-a-world is a condition that behaves classically with respect to negation. (.1) $\neg p$ is true at w if and only if it is not the case that p is true at w, and (.2) p is true at w if and only if it is not the case that $\neg p$ is true at w:

(.1) $w \models \neg p \equiv \neg w \models p$

(.2) $(w \models p) \equiv \neg w \models \neg p$

These facts show that negation is coherent with respect to *truth at* a possible world. Furthermore, we may universally generalize (.1) and (.2) with respect to p, and since we know $\varphi\downarrow$ and $\psi\downarrow$ for any formulas φ and ψ (104.2), we obtain theorem schemata by substituting φ for p in (.1) and (.2), keeping in mind the proviso for typed object theory mentioned in footnote 281.

(**530**) **Remark:** Situations, Worlds, Propositions, and Double Negation. Note that when we consider situations only (and not possible worlds), then $s \models q$ and $s \models \neg\neg q$ are not provably equivalent. To see this, recall that since s is a situation, $s \models q$ is equivalent, by theorem (471), to $s[\lambda y\, q]$, and $s \models \neg\neg q$ is equivalent to $s[\lambda y\, \neg\neg q]$. But since our theory doesn't imply that $[\lambda y\, \neg\neg q] = [\lambda y\, q]$, we cannot infer $s[\lambda y\, q]$ from $s[\lambda y\, \neg\neg q]$, or vice versa. Clearly, both $[\lambda y\, q]$ and $[\lambda y\, \neg\neg q]$ exist and it is provable that each necessarily implies the other, in the sense of (442.1). But situations aren't always 2-ary closed under necessary implication: from the facts that sF and $\Box\forall x(Fx \equiv Gx)$, it doesn't follow that sG. However, possible worlds are provably n-ary closed under necessary implication. So $w \models q$ is provably equivalent to $w \models \neg\neg q$.

(**531**) **Theorem:** There is a Unique Actual World. We first prove (.1) there exists a possible world that makes true all and only true propositions. Then we establish (.2) there is a unique actual world:

(.1) $\exists w \forall p(w \models p \equiv p)$

(.2) $\exists! w Actual(w)$

Since these theorem imply, *a fortiori*, that there are possible worlds, we know that the restricted quantifiers $\forall w$ and $\exists w$ behave classically in the sense that $\forall w \varphi \rightarrow \exists w \varphi$; cf. Remark (342).

Since both (.1) and (.2) are modally strict, it follow by RN that both are necessary truths. The necessitation of (.1), $\Box\exists w \forall p(w \models p \equiv p)$, intuitively asserts, in terms of semantically primitive possible worlds, that at every semantically primitive possible world $\boldsymbol{w}$, there is a object-theoretic possible world w' that encodes a proposition p if and only if p is true at $\boldsymbol{w}$. The necessitation of (.2), $\Box\exists! w Actual(w)$, intuitively asserts, in terms of semantically primitive possible worlds, that at every semantically primitive possible world $\boldsymbol{w}$, there is a object-theoretic possible world w' that is actual with respect to $\boldsymbol{w}$, i.e., by definition of *actual* (492), that every proposition w' encodes is true at $\boldsymbol{w}$.[285] So whereas theorem (517) is a universal claim that implies each possible world might have been actual, the proof of (.1) and (.2) actually gives a method for constructing, in any modal context, the object-theoretic possible world that encodes all and only the truths of that context.

After a short digression concerning another way in which the present theorem has significance, we'll work our ways towards a definition of a new term that refers to the unique object that is an actual world.

(**532**) **Remark:** On the Uniqueness of the Actual World. On the present theory, the claim that there is a unique actual world is provable, necessary, and known *a priori*. By contrast, on the concretist conception of *possible world*, the claim that there is a unique actual world appears to be a contingent fact, known only *a posteriori* and hence, not susceptible to proof. It may be that nothing like the previous theorem is available to one adopting the concretist conception of possible world.

Moreover, as far as I can tell, philosophers who've adopted a traditional abstractionist theory of possible worlds can offer a proof of the existence of a *unique* actual world only if they adopt the requirement that propositions are identical when necessarily equivalent, i.e., that $\Box(p \equiv q) \rightarrow p = q$. For if one allows that there are distinct, but necessarily equivalent propositions, then if possible worlds are identified as maximal and possible propositions (or states of affairs), one can prove the existence of multiply distinct actual worlds, at least given the following standard definition: p is a possible world iff (i) $\forall q((p \Rightarrow q) \vee (p \Rightarrow \neg q))$ and (ii) $\Diamond p$.[286] For suppose (a) p_1 is such a possible world, and (b) p_1 is actual, either in the sense that p_1 is true or in the sense that every proposition necessarily implied by p_1 is true. Then consider any proposition p_2 that is necessarily equivalent to p_1 (in the sense that $p_1 \Leftrightarrow p_2$) but *distinct* from p_1. Then it follows that p_1 and p_2 necessarily imply the same propositions.[287] So since p_1 satisfies clause (i), so does p_2 (exercise). Moreover,

[285] Of course, it is also true, and derivable, that at each semantically primitive possible world $\boldsymbol{w}$, there is a object-theoretic possible world w' that encodes a proposition p just in case *it is actually the case that* p. But this is not what these theorems assert. Once we form the description $\imath w Actual(w)$, its denotation will the object-theoretic possible world w such that it is actually the case that every proposition w makes true is true.

[286] Clause (i) is a maximality condition and clause (ii) guarantees consistency. And one could simplify the definition by stipulating instead: p is a possible world iff $\Diamond\forall q((p \Rightarrow q) \vee (p \Rightarrow \neg q))$. The objection, in what follows, to this conception of possible worlds doesn't apply to possible worlds conceived as sets of propositions. But then the conception of worlds as sets is not a *theory* of possible worlds but only a model of them, for reasons mentioned previously.

[287] The proof is a variant of left-to-right direction of (443.3). Assume $p_1 \Leftrightarrow p_2$. Then by (525), we know both:

(ϑ) $\Box(p_1 \rightarrow p_2)$

(ξ) $\Box(p_2 \rightarrow p_1)$

if p_1 is actual in the sense of being true, so is p_2, and if p_1 is actual in the sense that every proposition necessarily implied by p_1 is true, then so is p_2. So, there are two distinct actual worlds on this theory. The abstractionist can either derive the existence of a unique actual world by collapsing necessarily equivalent propositions or both allow for distinct but necessarily equivalent propositions and give up the claim that there is a unique actual world. Such a dilemma is not faced on the present theory. For a fuller discussion, see Zalta 1988 (73–4), 1993 (393–94), and McNamara 1993. A similar dilemma faces anyone who takes possible worlds to be properties without requiring the doubtful principle that $\Box\forall x(Fx \equiv Gx) \to F = G$.

(**533**) **Theorem:** A Significant Description. It follows immediately from (531.2) that the actual world exists:

$$\imath w Actual(w)\downarrow$$

The proof of this theorem is modally strict because it rests on (176.2), not on (144.2)★.

(**534**) **Definition:** The Actual World. Since we know that the actual world exists, we can name it:

$$\boldsymbol{w}_{\alpha} =_{df} \imath w Actual(w)$$

Note that the new constant symbol $\boldsymbol{w}_{\alpha}$ is a boldface, italic $\boldsymbol{w}$ decorated by a Greek α; this highlights the special character of the object denoted as the first among possible worlds. We do *not* use a plain italic w, for the resulting symbol would be conventionally used as a special variable whose value would depend on the value of w.

(**535**) **Theorems:** The True and The Actual World. Recall that we defined The True ($\top$) in (302.1) as the abstract object that encodes all and only those properties F of the form $[\lambda y\, p]$ for some true proposition p. It may come as a surprise that The True is identical with the actual world:

$$\top = \boldsymbol{w}_{\alpha}$$

We have therefore derived a fact about The True and the actual world suggested by Dummett in the following passage (1981, 180):

> If we take seriously Frege's manner of speaking in 'Uber Sinn und Bedeutung', the True must contain within itself the referents of the parts of all true sentences, and will admit a decomposition corresponding to each true

To show $\forall q(p_1 \Rightarrow q \equiv p_2 \Rightarrow q)$, it suffices by GEN to show $p_1 \Rightarrow q \equiv p_2 \Rightarrow q$. ($\to$) Suppose $p_1 \Rightarrow q$, i.e., $\Box(p_1 \to q)$. From (ξ) and this last result, it follows that $\Box(p_2 \to q)$, i.e., $p_2 \Rightarrow q$. ($\leftarrow$) Suppose $p_2 \Rightarrow q$, i.e., $\Box(p_2 \to q)$. then from (ϑ) and this last result, it follows that $\Box(p_1 \to q)$, i.e., $p_1 \Rightarrow q$.

> sentence. It thus becomes, in effect, an immensely complex structure, as it were the single all-inclusive Fact, which is how Kluge conceives of it, making it virtually indistinguishable from the world.

Dummett is referencing Kluge 1980 in this passage.

(**536**) ⋆**Theorem:** Truth and Truth At the Actual World. From (535), (305.1)⋆, (308.3), and the definition of *ExtensionOf*(x,p) (306), it follows that (.1) a proposition p is true if and only if the actual world $\boldsymbol{w}_\alpha$ is the extension of p:

(.1) $p \equiv \boldsymbol{w}_\alpha = \imath x ExtensionOf(x,p)$

This is an important and unheralded principle that relates true propositions, the actual world, and the extensions of propositions.

It also follows that (.2) a proposition p is true iff p is true at the actual world:

(.2) $p \equiv \boldsymbol{w}_\alpha \models p$

The proof shows that an earlier fact about The True (305.3)⋆, namely $p \equiv \top\Sigma p$, can be transformed to take on new significance. This theorem is another case where the non-modal strictness of the conditional results from the fact that the truth of one side varies from modal context to modal context, while the truth of the other side does not.

(**537**) **Theorem:** The Actual World Is a Possible World and Maximal. While it is trivial to prove that $\boldsymbol{w}_\alpha$ is both a possible world and maximal from (145.2)⋆ and (521), the claims in fact have modally strict proofs:

(.1) *PossibleWorld*$(\boldsymbol{w}_\alpha)$

(.2) *Maximal*$(\boldsymbol{w}_\alpha)$

By contrast, *Actual*$(\boldsymbol{w}_\alpha)$ isn't a modally strict theorem, although it is easy to prove it by *non*-modally strict means. However, we'll soon establish, by modally strict means, that $\boldsymbol{w}_\alpha$ is actual*, in the sense defined in Remark (494).

(**538**) **Theorem:** True at the Actual World and Actually True. The theorems in (537) allow us to prove a modally strict counterpart of (536.2)⋆, namely, that p is true at the actual world if and only if it is actually the case that p:

$$\boldsymbol{w}_\alpha \models p \equiv \mathcal{A}p$$

Note that the left-to-right direction of this theorem implies, by GEN, that $\forall p(\boldsymbol{w}_\alpha \models p \to \mathcal{A}p)$. So given the notion of *Actual** defined in Remark (494), it follows that *Actual**$(\boldsymbol{w}_\alpha)$ is a modally strict theorem.

(**539**) ⋆**Theorem:** Possible Worlds Other Than $\boldsymbol{w}_\alpha$ Aren't Actual. It follows immediately that possible worlds other than $\boldsymbol{w}_\alpha$ fail to be an actual world:

$$w \neq \boldsymbol{w}_\alpha \to \neg Actual(w)$$

Exercise. Develop a modally strict proof of $w \neq \boldsymbol{w}_\alpha \to \neg Actual^*(w)$.

(**540**) ★**Theorem:** Parts of the Actual World. Actual situations are part of the actual world:

$$Actual(s) \equiv s \trianglelefteq \boldsymbol{w}_\alpha$$

Cf. Barwise 1989a (261), where actual situations are *defined* to be the ones that are part of the actual world.

(**541**) ★**Theorems:** Facts About the Actual World. (.1) A proposition p is true at the actual world iff the actual world exemplifies being such that p. Moreover, (.2) a proposition p is true iff the proposition *the-actual-world-exemplifies-being-such-that-p* is true at the actual world:

(.1) $\boldsymbol{w}_\alpha \models p \equiv [\lambda y\ p]\boldsymbol{w}_\alpha$

(.2) $p \equiv \boldsymbol{w}_\alpha \models [\lambda y\ p]\boldsymbol{w}_\alpha$

(.2) is especially interesting because when p is true, we may derive from (.2) a statement of the form $\sigma \models \varphi(\sigma)$, where $\varphi(\sigma)$ is a formula in which the situation term σ occurs. In situation theory, statements of this form indicate that situation σ is *nonwellfounded*, since σ *makes true* a fact about itself. Note that when p is true, then (.2) implies that $[\lambda y\ p]\boldsymbol{w}_\alpha$ is true at $\boldsymbol{w}_\alpha$. And when $\neg p$ is true, then (.2) implies that $[\lambda y\ \neg p]\boldsymbol{w}_\alpha$ is true at $\boldsymbol{w}_\alpha$. So, in either case, $\boldsymbol{w}_\alpha$ is nonwellfounded in the above sense: a fact of the form $\varphi(\boldsymbol{w}_\alpha)$ is true at $\boldsymbol{w}_\alpha$. Thus, (.1) and (.2) decide Choices 8 (266) and 10 (268) in Barwise 1989a: situations can be constituents of facts and at least some situations are nonwellfounded.

(**542**) **Lemmas:** Some Basic Facts About Modality, Situations, Possible Worlds, and Truth At. (.1) If it is possible that p, then there might be a possible world at which p is true; (.2) if there might be a possible world at which p is true, then there is a possible world at which p is true; (.3) if p is true, then for all situations s, if all and only true propositions are true in s then p is true in s; (.4) if p is necessarily true, then it is necessary that for all situations s, if all and only true propositions are true in s, p is true in s; (.5) if necessarily every situation is such that φ, then every situation is necessarily such that φ; (.6) if p is true at every possible world w, then it is necessarily so; and (.7) if the claim that p is true at every possible world is necessary, then p is necessary:

(.1) $\Diamond p \to \Diamond \exists w(w \models p)$

(.2) $\Diamond \exists w(w \models p) \to \exists w(w \models p)$

(.3) $p \to \forall s(\forall q(s \models q \equiv q) \to s \models p)$

(.4) $\Box p \rightarrow \Box\forall s(\forall q(s \models q \equiv q) \rightarrow s \models p)$

(.5) $\Box\forall s\varphi \rightarrow \forall s\Box\varphi$

(.6) $\forall w(w \models p) \rightarrow \Box\forall w(w \models p)$

(.7) $\Box\forall w(w \models p) \rightarrow \Box p$

These lemmas are used to simplify the proofs of the fundamental theorems of possible world theory, to which we now turn. Note that (.5) establishes that the Converse Barcan Formula (167.2) holds when restricted to situations.

(**543**) **Theorems:** Fundamental Theorems of Possible World Theory. The foremost principles of possible world theory are that (.1) it is possible that p iff there is a possible world at which p is true; (.2) it is necessary that p iff p is true at all possible worlds; (.3) it is not possible that p iff there is no possible world at which p is true; and (.4) it is not necessary that p iff there is a possible world at which p fails to be true:

(.1) $\Diamond p \equiv \exists w(w \models p)$

(.2) $\Box p \equiv \forall w(w \models p)$

(.3) $\neg\Diamond p \equiv \neg\exists w(w \models p)$

(.4) $\neg\Box p \equiv \exists w\neg(w \models p)$

In the Appendix, there is a proof of (.1) that makes use of (542.1) and (542.2). In addition, there are two proofs of (.2), one that makes use of (542.4) – (542.7) and a simpler one that makes use of (.1). The proofs of (.3) and (.4) are left as simple exercises.

Note that (.1) is a strengthened version of the claim that every way that a world could possibly be is a way that some world is (Lewis 1986, 2, 71, 86). When we analyze *ways a possible world might be* as possibly true propositions, then Lewis's principle is the left-to-right direction of (.1). For a fuller discussion, see Menzel & Zalta 2014.

(**544**) **Theorem:** Facts about Modality and and the Existence of Possible Worlds: (.1) p is necessary if and only if there exists a possible world where p is necessary; (.2) p is necessary if and only if at every possible world, p is necessary; (.3) p is possible if and only if there exists a possible world where p is possible; and (.2) p is possible if and only if at every possible world, p is possible:

(.1) $\Box p \equiv \exists w(w \models \Box p)$

(.2) $\Box p \equiv \forall w(w \models \Box p)$

(.3) $\Diamond p \equiv \exists w(w \models \Diamond p)$

(.4) $\Diamond p \equiv \forall w(w \models \Diamond p)$

(**545**) **Theorem:** Truth At A World and the Connectives, Quantifiers, and Modal Operators. Since we already know how truth-at-a-world behaves with respect to negation (529), we begin by examining how it behaves with respect to the other connectives. (.1) $p \,\&\, q$ is true at a possible world w if and only if both p is true at w and q is true at w; (.2) $p \rightarrow q$ is true at a possible world iff if p is true at that possible world, then q is true at that possible world; (.3) $p \vee q$ is true at a possible world w if and only if either p is true at w or q is true at w; (.4) $p \equiv q$ is true at a possible world w if and only if: p is true at w iff q is true at w:

(.1) $w \models (p \,\&\, q) \equiv ((w \models p) \,\&\, (w \models q))$

(.2) $w \models (p \rightarrow q) \equiv ((w \models p) \rightarrow (w \models q))$

(.3) $w \models (p \vee q) \equiv ((w \models p) \vee (w \models q))$

(.4) $w \models (p \equiv q) \equiv ((w \models p) \equiv (w \models q))$

Clearly, we may universally generalize on the variables p and q in (.1) – (.4). So since $\varphi\downarrow$ and $\psi\downarrow$, for any formulas φ and ψ (104.2), the above hold for any formulas φ and ψ that may be uniformly substituted for p and q, respectively, keeping in mind the proviso for typed object theory mentioned in footnote 281.

Next, we show that the quantifers work similarly, except now it makes sense to formulate these claims, in the first instance, as schemata. (For the reasons, see footnote 283.) (.5) it is true at w that every α is such that φ if and only if every α is such that it is true at w that φ, and (.6) it is true at w that some α is such that φ if and only if some α is such that it is true at w that φ:

(.5) $(w \models \forall\alpha\varphi) \equiv \forall\alpha(w \models \varphi)$

(.6) $(w \models \exists\alpha\varphi) \equiv \exists\alpha(w \models \varphi)$

Now, switching back to propositional variables, it is easy to establish:[288]

(.7) $(w \models \Box p) \rightarrow \Box w \models p$

[288] If we use recent theorems to simplify the proof of (511.5), we obtain the following proof of (.8). Since there are contingently true propositions (217.1), let r be such, so that we know by (213.1) and facts about modality that:

(A) r

(B) $\neg\Box r$

By T$\Diamond$, (A) implies $\Diamond r$, and so a fundamental theorem of possible worlds (543.1) implies $\exists w(w \models r)$. Suppose w_1 is such a possible world, so that we know $w_1 \models r$. Then by (519.1), $\Box w_1 \models r$. So by &I and $\exists$I, it remains only to show $\neg w_1 \models \Box r$. Suppose, for reductio, $w_1 \models \Box r$. Then $\exists w(w \models \Box r)$. So by (544.1), $\Box r$, which contradicts (B).

(.8) $\exists w \exists p((\Box w \models p) \,\&\, \neg w \models \Box p)$

(.9) $(\Diamond w \models p) \rightarrow w \models \Diamond p$

(.10) $\exists w \exists p((w \models \Diamond p) \,\&\, \neg \Diamond w \models p)$

Exercise: In light of the Exercise at the end of (511), determine whether there is a proof of the following:

$$(w \models \mathcal{A}p) \equiv \mathcal{A}w \models p$$

If there is a proof, indicate whether or not it is modally strict. If you can't find a proof of the biconditional, can you find a proof of at least one direction?

(**546**) **Remark:** Reconciling Lewis Worlds with Possible Worlds. There is a way to reconcile, at least to some extent, the present theory of possible worlds with Lewis's conception of possible worlds. Note that the present theory of possible worlds doesn't tell us which ordinary objects are concrete at which worlds. But we may define:

> x is a *physical universe at* w if and only if (a) x is concrete at w, (b) every object y that is concrete at w is a part of x, and (c) x exemplifies exactly the propositional properties $[\lambda y\, p]$ such that p is true at w.

If given a part-whole relation $\sqsubseteq$ for concrete objects, one could formalize this definition as follows:[289]

$$\begin{aligned}&\mathit{PhysicalUniverseAt}(x,w) =_{df}\\ &\quad w \models E!x \;\;\&\;\; \forall y(w \models E!y \rightarrow y \sqsubseteq x) \;\;\&\;\; \forall p([\lambda y\, p]x \equiv w \models p)\end{aligned}$$

Given this definition and the assumption that $\sqsubseteq$ obeys a standard mereological principle, namely, that for any two objects x, y that are concrete at w there is an object z that is concrete at w of which x and y are a part, one might reasonably assert that for each possible world w at which there are concrete objects, there is a universe at w, i.e.,

$$\forall w(w \models \exists x E!x \rightarrow \exists y \mathit{PhysicalUniverseAt}(y,w))$$

[289] It is important to understand why the third conjunct is included in the definition of $\mathit{PhysicalUniverseAt}(x,w)$. Without this clause, a physical universe at w would be defined solely in terms of the objects that are concrete at w. Consider the case in which there are two possible worlds w_1 and w_2 such that (i) exactly the same objects are concrete at w_1 and w_2 and (ii) different propositions are true at w_1 and w_2. Thus, as we know, clause (ii) implies $w_1 \neq w_2$. But without the third clause in the definition, any physical universe at w_1 would be identified with any physical universe at w_2, since they have the same concrete parts. Of course, Lewis doesn't worry about such a result, since individuals are all world bound; his view doesn't allow for different possible worlds with exactly the same individuals. But it would be a problem for the above reconstruction of Lewis worlds without the third clause in the definition.

Furthermore if $\sqsubseteq$ obeys the principle that for each w there is at most one concrete object at w for which every concrete object at w is a part, then it would follow that for each possible world w at which there are concrete objects, there is a unique universe at w, i.e.,

$$\forall w(w \models \exists x E!x \rightarrow \exists! y PhysicalUniverseAt(y, w))$$

Indeed, from these claims, one should be able to prove that if there is a physical universe for w, there is a unique physical universe for w, i.e.,

$$\exists x PhysicalUniverseAt(x, w) \rightarrow \exists! x PhysicalUniverseAt(x, w)$$

Now if one were to analyze Lewis's notion of a possible world as a physical universe, then principles akin to those that govern his theory of possible worlds would govern universes. For it would follow, if one assumes the existence of concrete objects, that there is a universe for the actual world, and that universes at other possible worlds are no different in kind from the universe of the actual world – all universes are ordinary objects. Whereas Lewis holds that there are universes just like our own but which are different mereological sums of concrete objects, the above analysis yields that there are universes just like our own but which are different mereological sums of ordinary objects. This reconciles, at least to some extent, the Lewis conception of possible worlds with the present one.

(**547**) **Theorems:**. Existence of Non-Actual Possible Worlds. The Fundamental Theorems imply: (.1) if some proposition is contingently true, then there exists a nonactual possible world; (.2) if some proposition is contingently false, then there exists a nonactual possible world:

(.1) $\exists p(ContingentlyTrue(p)) \rightarrow \exists w(\neg Actual(w))$

(.2) $\exists p(ContingentlyFalse(p)) \rightarrow \exists w(\neg Actual(w))$

So the existence of nonactual possible worlds depends upon there being a contingent truth or a contingent falsehood.

From (.1) and (217.1) ("there are contingent truths"), or from (.2) and (217.2) ("there are contingent falsehoods"), it immediately follows that (.3) there are nonactual possible worlds, and hence, that (.4) there are at least two possible worlds:

(.3) $\exists w \neg Actual(w)$

(.4) $\exists w \exists w'(w \neq w')$

(**548**) **Remark:** Nonactual Possible Worlds. These results show not only that our theory implies a claim stronger than Alternative 1.2 ("There may be more than one world") in Barwise 1989a (260), but also that those philosophers who have claimed that there are nonactual possible worlds are provably correct. Note that axiom (45.4) plays a key role, since the proofs of (547.3) and (547.4) rest on (217.1), which in turn depends on (45.4). Interestingly, if a Humean were to reject axiom (45.4), they could nevertheless accept an interesting fragment of the above theory of possible worlds. Without axiom (45.4), one can't derive the existence of nonactual possible worlds, though the Fundamental Theorems (543.1) – (543.4) remain derivable and true (see Menzel & Zalta 2014). Given the many theorems in which (45.4) has proved its worth, theory comparison is in order should one balk at its acceptance.

(**549**) **Theorem:** Derivation of the Kaplan Axiom (Contributed by Daniel West). In a one-page abstract of a talk, Kaplan (1970) identified a number of axioms for S5 modal logic extended with quantifiers ranging over propositional variables. One of the axioms stands out, namely, the existence claim: there is a true proposition that necessarily implies every true proposition. Daniel West found a proof of this claim and has kindly granted permission to reproduce it here:

$$\exists p(p \,\&\, \forall q(q \to \Box(p \to q)))$$

Hint: The witness that West found is a complex proposition that is not describable purely in terms of exemplification formulas. But it can be shown, by modally strict means, that the proposition in question is true and necessarily implies every truth.

(**550**) **Remark:** Iterated Modalities. It is worth spending some time showing that we can derive the correct possible worlds analysis of *iterated* modalities within the language of object theory itself. Let us represent the fact that *Obama might have had a son who might have become president* as:

(φ) $\Diamond\exists x(Sxo \,\&\, \Diamond Px)$

We now proceed to show that object theory yields the following possible worlds analysis of φ, namely:

(ψ) $\exists w\exists x((w \models Sxo) \,\&\, \exists w'(w' \models Px))$

This asserts that for some possible world w, there is an object x such that (a) x *is a son of Obama* at w and (b) at some possible world w', x *is president*.

To establish that (ψ) follows from (φ), note that by BF$\Diamond$ (167.3), (φ) is equivalent to: there is an object x such that, possibly, both x is Obama's son and possibly x is president, i.e.,

$$\exists x\Diamond(Sxo \,\&\, \Diamond Px)$$

Suppose b is such an object, so that we know:

$$\Diamond(Sbo \mathbin{\&} \Diamond Pb)$$

This is equivalent, by Fundamental Theorem (543.1), to:

$$\exists w(w \models (Sbo \mathbin{\&} \Diamond Pb))$$

Suppose w_1 is such a world, so that we know:

$$w_1 \models (Sbo \mathbin{\&} \Diamond Pb)$$

By (545.1), we know that this last fact implies and, indeed, is equivalent to:

$$(\vartheta)\quad (w_1 \models Sbo) \mathbin{\&} (w_1 \models \Diamond Pb)$$

Now it important to recognize that to analyze the modal operator in the right conjunct of (ϑ) in terms of possible worlds, we may *not* immediately appeal to an instance of Fundamental Theorem (543.1) to substitute $\exists w(w \models Pb)$ for $\Diamond Pb$ in the right conjunct of (ϑ).[290] Instead, we next infer $\exists w(w \models \Diamond Pb)$ from the second conjunct of (ϑ). Then, by Fundamental Theorem (543.1), it follows that $\Diamond\Diamond Pb$, which by (165.8) is equivalent to $\Diamond Pb$. So, again, by Fundamental Theorem (543.1), we know that $\exists w'(w' \models Pb)$. So let us put together this last fact with the first conjunct of (ϑ), to obtain:

$$(w_1 \models Sbo) \mathbin{\&} \exists w'(w' \models Pb)$$

From this it follows that:

$$\exists x((w_1 \models Sxo) \mathbin{\&} \exists w'(w' \models Px))$$

And by a second application of $\exists$I, we obtain:

$$(\psi)\quad \exists w \exists x((w \models Sxo) \mathbin{\&} \exists w'(w' \models Px))$$

Thus, in our system, (φ) implies its classic, possible-worlds truth conditions.

Of course, before it is fair to call (ψ) *truth conditions* for (φ), we have to show that (ψ) is equivalent to (φ). It remains, therefore, to derive (φ) from (ψ). Here is a sketch of a modally strict proof:

[290] The reason we cannot use a Rule of Substitution to substitute $\exists w(w \models Pb)$ for $\Diamond Pb$ on the basis of the fact that $\vdash_\Box \Diamond Pb \equiv \exists w(w \models Pb)$ is that $\Diamond Pb$ is not a subformula of $w_1 \models \Diamond Pb$. If we expand the latter by its definition to $w_1[\lambda y\, \Diamond Pb]$, then it becomes clear why $\Diamond Pb$ is not one of its subformulas.

(1)	$\exists w \exists x((w \models Sxo) \,\&\, \exists w'(w' \models Px))$	(ψ)
(2)	$(w_2 \models Sco) \,\&\, \exists w'(w' \models Pc)$	Premise for $\exists$E, w_2 and c arbitrary
(3)	$(w_2 \models Sco) \,\&\, \Diamond Pc$	From (2), by Fund. Thm. (543.1)
(4)	$\Box\Diamond Pc$	From 2nd conjunct of (3), by (45.3)
(5)	$\forall w(w \models \Diamond Pc)$	From (4), by (543.2)
(6)	$w_2 \models \Diamond Pc$	From (5), by $\forall$E
(7)	$w_2 \models (Sco \,\&\, \Diamond Pc)$	From 1st conjunct (2), (6), by (545.1)
(8)	$\exists w(w \models (Sco \,\&\, \Diamond Pc))$	From (7), by $\exists$I
(9)	$\Diamond(Sco \,\&\, \Diamond Pc)$	From (8), by (543.1)
(10)	$\exists x \Diamond(Sxo \,\&\, \Diamond Px)$	From (9), by $\exists$I
(11)	(φ)	From (10), by CBF$\Diamond$ (167.4)

This example therefore shows us how to use the Fundamental Theorems to interderive, within object theory, ordinary modal propositions with iterated modalities and their classical possible-world truth conditions.[291] It is not too far off the mark to suggest that in addition to the Fundamental Theorems, the keys to the equivalence of (φ) and (ψ) are, for the left-to-right direction, the Barcan Formula ($\forall\alpha\Box\varphi \to \Box\forall\alpha\varphi$) and S4$\Diamond$ ($\Diamond\Diamond\varphi \to \Diamond\varphi$), and for the right-to-left direction, the 5 schema ($\Diamond\varphi \to \Box\Diamond\varphi$) and the Converse Barcan Formula ($\Box\forall\alpha\varphi \to \forall\alpha\Box\varphi$).

(**551**) **Theorem:** A Useful Equivalence Concerning Worlds and Objects. The following will prove to be a useful consequence of our theory of possible worlds when we investigate Leibniz's modal metaphysics: a proposition p is true at a

[291]The example discussed in this Remark was chosen because it has a form that is relevantly similar to an example that has figured prominently in the literature. In McMichael 1983 (54), we find:

> Consider the sentence:
>
> (5) It is possible that there be a person X who does not exist in the actual world, and who performs some action Y, but who might not have performed Y.
>
> This sentence is surely true. For example, John F. Kennedy could (logically) have had a second son who becomes a Senator, although he might have chosen to become an astronaut instead.

Clearly, we can simplify McMichael's example to:

> John F. Kennedy could have had a second son who becomes a Senator but might not have.

Then where Sxy represents x is a son of y, $S'x$ represents x is a senator, and k represents Kennedy, we could represent this example as:

$$\Diamond\exists x(Sxk \,\&\, S'x \,\&\, \Diamond\neg S'x)$$

The example discussed in our Remark further simplifies McMichael's example in two ways: (1) we have eliminated the conjunct $S'x$, and (2) we've made the embedded modal claim into a positive statement instead of a negative one. Thus, it should be clear that our simplifications of the example have not changed its essential features. If one represents McMichael's original example without the simplifications, the proof that the modal claim and its possible-worlds truth conditions are equivalent goes basically the same way, though the details do become more complex.

world w iff the proposition *that x exemplifies being such that p* is true at w:

$$w \models p \equiv w \models [\lambda y\, p]x$$

This theorem plays a role in the development of the theory of Leibnizian concepts in Chapter 13.

(**552**) **Theorem:** Possible Worlds and *Ex Contradictione Quodlibet*. We now prove a few theorems about possible worlds that will stand in contrast to theorems in Section 12.4 about impossible worlds. Note that it follows immediately from (518) by GEN that every possible world is consistent. By applying definitions and quantification theory to this universal claim, it can be transformed into the equivalent claim that there is no possible world w and proposition p such that p and $\neg p$ are both true at w. It also follows that there is no possible world w and proposition p such that $p \,\&\, \neg p$ is true at w:

(.1) $\neg\exists w \exists p(w \models (p \,\&\, \neg p))$

If we recall Remark (499), where we discussed the notion of *Consistency** and showed that it is independent of the notion of *Consistency*, then the above theorem can easily be transformed into the claim: every world is *consistent**.

This theorem has interesting consequences related to the traditional logical law *ex contradictione quodlibet*, which is almost always formulated in the formal mode as: every formula ψ is derivable from a contradiction of the form $\varphi \,\&\, \neg\varphi$, i.e., $\varphi \,\&\, \neg\varphi \vdash \psi$. This formal principle clearly governs our system.[292] However, when we formulate *ex contradictione quodlibet* in the material mode so that it applies to propositions, it becomes the easily-established theorem $(p \,\&\, \neg p) \Rightarrow q$. Hence by the fact that possible worlds are 2-modally closed and the definition of n-modal closure, it follows that (.2) if some proposition p is such that $p \,\&\, \neg p$ is true at possible world w, then every proposition is true at w:

(.2) $\exists p(w \models (p \,\&\, \neg p)) \rightarrow \forall q(w \models q)$

(.2) easily follows from the fact that negation of its antecedent is implied by (.1). Since every formula signifies a proposition, (.2) immediately yields the schema $w \models (\varphi \,\&\, \neg\varphi)) \rightarrow w \models \psi$.

(**553**) **Theorem:** Disjunctive Syllogism Holds at a Possible World. (.1) If $p \vee q$ is true at w and $\neg p$ is true at w, then q is true at w:

$$(w \models (p \vee q) \,\&\, (w \models \neg p)) \rightarrow w \models q$$

Since every formula is significant, this theorem implies that the following schema holds: $(w \models (\varphi \vee \psi) \,\&\, (w \models \neg\varphi)) \rightarrow w \models \psi$. Thus, disjunctive syllogism

[292] From the premise $\varphi \,\&\, \neg\varphi$, we may infer both φ and $\neg\varphi$, by &E. From φ, we may infer $\varphi \vee \psi$, for any ψ, by $\vee$I. But from $\varphi \vee \psi$ and $\neg\varphi$ it follows that ψ, by disjunctive syllogism (86.4.b).

holds both as an object-theoretic claim about propositions and worlds and as a metatheoretic schema about formulas and worlds.

(**554**) **Remark:** Final Observations on the Theory of Possible Worlds. The foregoing theory of possible worlds requires none of Leibniz's theological doctrines, such as his claims about what goes on in God's mind or his theodicy to explain the existence of evil. Indeed, I think Stalnaker's severe judgment (1976, 65), which we quoted at the beginning of this section, should be reevaluated in light of the foregoing analysis. Leibniz's structural vision about the space of worlds is no mere metaphor. It has been reconstructed as a scientifically and mathematically precise theory, and grounds the different analyses invoking possible worlds which Stalnaker developed in his own work.

As we've seen, worlds are objects that can be abstracted from ordinary predication and possibility, and it doesn't matter whether we take the locus of predication and possibility to be the physical world, our minds, or language. One need only accept that there is in fact a corpus of ordinary predications and possibilities, that within this corpus there are special patterns of propositions, and that possible worlds are nothing more than those patterns.

As abstractions, possible worlds have an intrinsic nature, defined by their encoded properties. As noted previously, the propositions that are true at world w characterize w. That is, if $w \models p$, then w *is* such that p, in the sense that w encodes $[\lambda y\, p]$. Possible worlds, as we've defined them, don't model or represent anything, despite being abstract; they are not *ersatz* worlds (Lewis 1986, 136ff). Possible worlds just *are* abstract logical objects characterized by the propositions true at them.[293]

Finally, note that once we extend our system by adding particular contingent truths,[294] we do not require any special, further evidence for believing in the existence of each nonactual possible world implied by (543.1). Epistemologically, we don't have to justify our knowledge of each possible world, e.g., by identifying some information pathway from each world back to us that explains and justifies our belief in its existence. Instead, given contingently true propositions as data, we can cite (543.1) as the principle that guarantees the existence of the relevant nonactual worlds and thereby justify our belief in

[293] It appears as though Lewis (1986, Chapter 3) would regard Wittgensteinian possible worlds, referenced in Wittgenstein 1921 (7, Propositions 1 and 1.1) in such claims as "The world is all that is the case" and "The world is the totality of facts, not of things", as ersatz. But I'm not sure why possible worlds characterized in terms of the propositions true at them should be labeled *representations* and thereby *ersatz*. The possible worlds characerized in the present text are *not* ersatz, given the facts just noted in the text. It bears emphasizing that encoding is a mode of predication, and so when $w \models p$ and *being such that p* is thereby predicated of w, the property in question *characterizes* w. So w *is* such that p, and doesn't represent something that is such that p.

[294] I am referring here to such claims such as: Obama doesn't have a son but might have, Obama has two daughters but might not have, there aren't million carat diamonds but there might have been, etc.

them. In turn, the justification of (543.1) rests on the fact that it is derivable from a very general theory with both inferential and explanatory power. Moreover, the justification also goes in the other direction: the theory itself receives justification from the fact that the above principles governing possible worlds are derivable from it.

12.3 World-Indexed Logical Objects and Relations

12.3.1 World-Indexed Truth-Values

In the next numbered item, we revise and enhance definition *truth-value of p* (286), in two ways: (1) by relativizing it to possible worlds, and (2) by taking advantage of the fact that truth-values are situations. We could relativize definition (286) to possible worlds and accomplish (1) by defining:

$$TruthValueAtOf(x,w,p) \equiv_{df} A!x \,\&\, \forall F(xF \equiv \exists q(w \models (q \equiv p) \,\&\, F = [\lambda y\, q]))$$

This would define a truth-value-at-w of p to be an abstract object x that encodes all and only the propositional properties F constructed out of propositions q that are materially equivalent to p at w. However, we can now take advantage of the fact that truth-values are situations (468) and accomplish (2) by defining a truth-value-at-w of p in a way that makes use of our simplified comprehension conditions for situations (486.1). A truth-value-at-w of p can be defined as a situation s that makes true all and only the propositions equivalent to p at w. We may formalize this as follows.

(**555**) **Definition:** Truth-Value At w Of p. We define: s is a *truth-value-at-w of p* if and only if s is a situation that makes true all and only the propositions q such that, at w, q is materially equivalent to p:

$$TruthValueAtOf(s,w,p) \equiv_{df} \forall q(s \models q \equiv w \models (q \equiv p))$$

We continue to use the rigid restricted variables $s, s', \ldots$ for situations and $w, w', \ldots$ for possible worlds.

(**556**) **Theorems:** Unique Existence of Truth-Values-at-Worlds of Propositions. It now follows that: (.1) there is a unique truth-value-at-w of p, and (.2) the truth-value-at-w of p exists:

(.1) $\exists! s\, TruthValueAtOf(s,w,p)$

(.2) $\imath s\, TruthValueAtOf(s,w,p)\downarrow$

Recall that since these theorems involve the free restricted variable w, they are implicitly conditionals. Though since we've proved the existence of possible worlds, we can derive unconditional existence claims.

(**557**) **Definition:** Notation for The Truth-Value-At-w Of p. Given (556.2) and our conventions for definitions, we may introduce notation to designate the truth-value-at-w of p:

$$\circ_w p \;=_{df}\; \imath s TruthValueAtOf(s,w,p)$$

This introduces $\circ_w p$ as a binary functional term with free variable p and free restricted variable w. Hence, given our conventions for definitions and restricted variables (339) an expression of the form $\circ_\kappa \Pi$ is significant when Π is any 0-ary term in which s isn't free and κ is known to be a possible world, either by proof or by hypothesis.[295]

(**558**) **Theorems:** Strict Canonicity of $\circ_w p$. In the usual way, it follows that $\circ_w p$ is (identical to) a canonical situation, as defined in (486):

(.1) $\circ_w p = \imath s \forall q(s \models q \equiv w \models (q \equiv p))$

Now if we let φ be $w \models (q \equiv p)$, then it follows that φ is a rigid condition on propositions, as this was defined in (260.1), for the following is established by modally strict means:

(.2) $\forall q((w \models (q \equiv p)) \to \Box w \models (q \equiv p))$

So $\circ_w p$ is strictly canonical, by (260.2), and is subject to theorem (261.2). It is then easy to establish, as a modally strict theorem, that (.3) $\circ_w p$ is a a situation that makes true exactly the propositions equivalent to p at w:

(.3) $\forall q(\circ_w p \models q \equiv w \models (q \equiv p))$

(.4) $TruthValueAtOf(\circ_w p, w, p)$

It follows relatively quickly from (.3), the fact that $\Box(p \equiv p)$, and a fundamental theorem of world theory that (.5) the truth-value-at-w of p makes p true:

(.5) $\circ_w p \models p$

Finally, the Fregean biconditional, (.6) the truth-value-at-w of p is identical to the truth-value-at-w of q if and only if it is true at w that p and q are materially equivalent:

(.6) $\circ_w p = \circ_w q \;\equiv\; w \models (p \equiv q)$

[295] This is explained by the fact that the above definition abbreviates:

$$\circ_y p \;=_{df}\; \imath x(PossibleWorld(y) \,\&\, TruthValueAtOf(s,y,p))$$

So if y is replaced by a term κ that fails to denote a possible world, then $PossibleWorld(\kappa)$ fails and so no situation s satisfies the condition $PossibleWorld(y) \,\&\, TruthValueAtOf(s,\kappa,p)$. In that case, the definiens would fail to be significant and, hence, so would the definiendum, by the Rule of Definition by Identity (73).

So whereas the Fregean biconditional for truth-values (300)★ is not a modally strict theorem, the analogous principle for world-indexed truth-values is.

(**559**) **Definitions:** The-True-at-w and The-False-at-w. We define: (.1) The True-at-w ($\top_w$) to be the situation that makes true exactly the propositions that are true at w; (.2) The False-at-w ($\bot_w$) to be the situation that makes true exactly the propositions that are false at w:

(.1) $\top_w =_{df} \imath s\forall p(s \models p \equiv w \models p)$

(.2) $\bot_w =_{df} \imath s\forall p(s \models p \equiv w \models \neg p)$

Given the free restricted variable w in these definitions, we may regard $\top_\kappa$ and $\bot_\kappa$ as significant only when κ is known to be a possible world, by proof or by hypothesis.

(**560**) **Theorems:** Strict Canonicity of The-True-at-w and The-False-at-w. Clearly, $\top_w$ and $\bot_w$ are (identical to) strictly canonical situations, as defined in (486), since the following modally strict theorems imply (.1) $w \models p$ is a rigid condition on propositions, and (.2) $w \models \neg p$ is rigid condition on propositions:

(.1) $\forall p((w \models p) \rightarrow \Box(w \models p))$

(.2) $\forall p((w \models \neg p) \rightarrow \Box(w \models \neg p))$

So $\top_w$ and $\bot_w$ are subject to theorem (261.2) and can be instantiated into their own defining descriptions by modally strict proofs. Thus, we have (.3) The-True-at-w makes true all and only the truths at w, and (.4) The-False-at-w makes true all and only the falsehoods at w:

(.3) $\forall p(\top_w \models p \equiv w \models p)$

(.4) $\forall p(\bot_w \models p \equiv w \models \neg p)$

It follows from (.3) and (474) that The-True-at-w just is w:

(.5) $\top_w = w$

(.5) is a more general result than (535), which asserts that $\top = \boldsymbol{w}_\alpha$.

(**561**) **Theorems:** Truth at w, The Truth-Value-at-w of p, and The True at w. (.1) p is true at w if and only if the truth-value-at-w of p is The-True-at-w; and (.2) p is false at w if and only if the truth-value-at-w of p is The-False-at-w.

(.1) $w \models p \equiv \circ_w p = \top_w$

(.2) $w \models \neg p \equiv \circ_w p = \bot_w$

(**562**) **Theorems:** World-Indexed Truth-Values and Modalities. It now follows that: (.1) p is necessary if and only if for every world w, the truth-value-at-w of p is (identical to) The-True-at-w; (.2) p is necessarily false if and only if for every world w, the truth-value-at-w of p is The-False-at-w; (.3) p is possible if and only if for some world w, the truth-value-at-w of p is The-True-at-w; and (.4) p is possibly false if and only if for some world w, the truth-value-at-w of p is The-False-at-w:

(.1) $\Box p \equiv \forall w(\circ_w p = \top_w)$

(.2) $\Box\neg p \equiv \forall w(\circ_w p = \bot_w)$

(.3) $\Diamond p \equiv \exists w(\circ_w p = \top_w)$

(.4) $\Diamond\neg p \equiv \exists w(\circ_w p = \bot_w)$

12.3.2 World-Indexed Extensions

(**563**) **Definitions:** An Extension at a World of a Property. We may now relativize definition (312.1) as follows: x is an extension-at-w of G if and only if x is abstract, G exists, and x encodes just those properties F such that at w, F is materially equivalent to G:

$$\mathit{ExtensionAtOf}(x, w, G) \equiv_{df} A!x \,\&\, G{\downarrow} \,\&\, \forall F(xF \equiv w \models \forall y(Fy \equiv Gy))$$

It is important here *not* to think of an extension at w of G as an extension of the property G-at-w. We'll *later* introduce such world-indexed properties as G-at-w (G_w) later and it will become clear, if only as an exercise, why the extension of the property G_w (namely, ϵG_w), which encodes the properties F that in fact are materially equivalent to G_w is not an extension-at-w of G ($\epsilon_w G$), which encodes the properties F that are materially equivalent to G at w. See Exercise in Remark (575).

(**564**) **Theorem:** The Extension at w of a Property Exists. It now follows that (.1) there exists a unique extension-at-w of G, and hence, that (.2) *the* extension-at-w of G exists:

(.1) $\exists! x \mathit{ExtensionAtOf}(x, w, G)$

(.2) $\imath x \mathit{ExtensionAtOf}(x, w, G){\downarrow}$

Since both w and G are free variables in these theorems, it follows by GEN that the latter hold for all properties and worlds.

(**565**) **Definition:** Notation for the Extension-at-w of G. We may therefore introduce notation for the extension-at-w of G as follows:

$$\epsilon_w G \ =_{df} \ \imath x ExtensionAtOf(x, w, G)$$

Since the existence of G is built into the definition of $ExtensionAtOf(x, w, G)$ and this definition uses the restricted variable w, expressions of the form $\epsilon_\kappa \Pi$ are binary functional terms that are significant only when it is known, by proof or by hypothesis, that κ is a possible world and Π is a significant unary relation term.

(**566**) **Theorems:** Strict Canonicity of $\epsilon_w G$. By now well-established reasoning, (.1) $\epsilon_w G$ is (identical to) a canonical object:

(.1) $\epsilon_w G = \imath x(A!x \,\&\, \forall F(xF \equiv w \models \forall y(Fy \equiv Gy)))$

If we now let φ be $w \models \forall y(Fy \equiv Gy)$, then it follows that (.2) φ is a rigid condition on properties, as this was defined in (260.1):

(.2) $\forall F(w \models \forall y(Fy \equiv Gy) \rightarrow \Box w \models \forall y(Fy \equiv Gy))$

So we know that $\epsilon_w G$ is strictly canonical, by (260.2). Hence, it is subject to theorem (261.2). It is therefore easy to establish, by modally strict means, that (.3) $\epsilon_w G$ is an abstract object that encodes exactly those properties F such that at w, F is materially equivalent to G, and (.4) $\epsilon_w G$ is an extension-at-w of G:

(.3) $A!\epsilon_w G \,\&\, \forall F(\epsilon_w GF \equiv w \models \forall y(Fy \equiv Gy))$

(.4) $ExtensionAtOf(\epsilon_w G, w, G)$

Finally, it follows relatively quickly from the second conjunct of (.3) that (.5) the extension-at-w of G encodes G:

(.5) $\epsilon_w GG$

This is an encoding claim in which the individual term is the restricted functional term $\epsilon_w G$.

(**567**) **Theorem:** World-Indexed Pre-Law V and World-Indexed Law V. It now follows that (.1) if x is an extension-at-w of G and y is an extension-at-w of H, then $x = y$ if and only if at w, G and H are materially equivalent; and (.2) the extension-at-w of F is identical to the extension-at-w of G if and only if, at w, F and G are materially equivalent:

(.1) $(ExtensionAtOf(x, w, G) \,\&\, ExtensionAtOf(y, w, H)) \rightarrow$
$(x = y \equiv w \models \forall z(Gz \equiv Hz))$

(.2) $\epsilon_w F = \epsilon_w G \ \equiv \ w \models \forall z(Fz \equiv Gz)$

So though Frege's Basic Law V (328)$\star$ is not modally-strict, its world-indexed version is.

(**568**) **Remark:** Suggestions for Further Research. Define: x is, *at* w, a *class of* Gs if and only if x is an extension-at-w of G. Then define world-indexed membership: y is, *at* w, an *element of* x iff x is, *at* w, a class of Gs and, at w, Gy:

$$y \in_w x \ \equiv_{df} \ \exists G(ClassOfAt(x,w,G) \,\&\, w \models Gy)$$

Though we shall not develop these ideas further, the reader should consider formulating and proving some interesting claims regarding $y \in_w x$, as well as defining further notions (such as *the* class, at w, of Gs) and proving facts about them.

12.3.3 World-Indexed Relations

The key results in this section, namely, the proof of the existence of world-indexed relations and the proof that every relation has a rigidification, are consequences of Kirchner's Theorem and were first reported by Daniel Kirchner in personal communications. See footnotes 296 and 299 below.

(**569**) **Theorems:** Relations Defined in Terms of Truth at a Possible World. We begin by establishing that (.1) if *being an x such that φ* exists, then *being an x such that, at w, φ is true* exists:[296]

(.1) $[\lambda x\, \varphi]{\downarrow} \rightarrow [\lambda x\, w \models \varphi]{\downarrow}$

$[\lambda x\, w \models \varphi]$ isn't yet what we shall call a world-indexed relation. As (.1) shows, this relation exists only conditionally. We'll define *world-indexed* relations in (570.1) below; these will exist unconditionally for every relation (570.2). Note that (.1) generalizes to:

(.2) $[\lambda x_1 \ldots x_n\, \varphi]{\downarrow} \rightarrow [\lambda x_1 \ldots x_n\, w \models \varphi]{\downarrow}$ $\qquad (n \geq 0)$

In constrast to the relations defined in (.1) and (.2), we now single-out a special class of relations that exist axiomatically and unconditionally, as instances of (39.2), namely, *being an $x_1,\ldots,x_n$ such that, at w, $x_1,\ldots,x_n$ exemplifies F*:

(.3) $[\lambda x_1 \ldots x_n\, w \models Fx_1 \ldots x_n]{\downarrow}$ $\qquad (n \geq 0)$

These are instances of (39.2) because $[\lambda x_1 \ldots x_n\, w \models Fx_1 \ldots x_n]$ is a core λ-expression, as this was defined in (9.2): no variable bound by the λ occurs in encoding position in the matrix $w \models Fx_1 \ldots x_n$, even when we consider the definientia of the defined terms in $w \models Fx_1 \ldots x_n$.[297]

(**570**) **Definition and Theorem:** World-Indexed Relations. We now define the world-indexed relation F^n_w ('F^n-*at*-w') as $[\lambda x_1 \ldots x_n\, w \models F^n x_1 \ldots x_n]$:

[296] I'm indebted to Daniel Kirchner for noting (personal communication: 5 March 2020) this important consequence of the theorem he contributed, which was reported in (271).

[297] The matrix $w \models Fx_1 \ldots x_n$ is defined as (470):

$$Situation(w) \,\&\, w\Sigma Fx_1 \ldots x_n$$

(.1) $F^n_w =_{df} [\lambda x_1 \ldots x_n\, w \models F^n x_1 \ldots x_n]$ $(n \geq 0)$

Thus, it follows that (.2) for every n-ary relation F and possible world w, *being F at w* exists:

(.2) $\forall F^n \forall w (F^n_w\downarrow)$ $(n \geq 0)$

We have therefore proved, in the object language, the existence of *world-indexed relations among individuals*, and have no need to stipulate their existence in the metalanguage. This stands in contrast to Williamson 2013 (237), where world-indexed relations of *every* higher type are stipulated to exist. We'll prove the existence of world-indexed relations of every higher type in Chapter 15, in (972.2) and (972.3).

(**571**) **Definition:** Rigid and Rigidifying Relations. To formulate and prove an important group of theorems, let us say that a relation F^n is *rigid* if and only if necessarily, for any objects $x_1,\ldots,x_n$, if $x_1 \ldots x_n$ exemplify F^n, then necessarily $x_1,\ldots,x_n$ exemplify F^n:

(.1) $Rigid(F^n) \equiv_{df} F^n\downarrow \,\&\, \Box \forall x_1 \ldots \forall x_n (F^n x_1 \ldots x_n \rightarrow \Box F^n x_1 \ldots x_n)$ $(n \geq 0)$

In the 0-ary case, the definition stipulates that $Rigid(p) \equiv_{df} p\downarrow \,\&\, \Box(p \rightarrow \Box p)$.

Moreover, where $n \geq 0$, let us say that F^n *rigidifies* G^n just in case F^n is rigid and is exemplified by exactly the same objects as G^n:

(.2) $Rigidifies(F^n, G^n) \equiv_{df} Rigid(F^n) \,\&\, \forall x_1 \ldots \forall x_n (F^n x_1 \ldots x_n \equiv G^n x_1 \ldots x_n)$

In the 0-ary case, $Rigidifies(p,q)$ just in case $Rigid(p)$ and p is materially equivalent to q.

(**572**) **Remark:** Digression on the Notions of Rigidity. Since we've now used the term 'rigid' and its cognates in various senses, it might serve well to review the various notions introduced thus far, to forestall confusion in what follows. We first describe how we've used the most recognizable notion of rigidity:

Expanding the definition of $Situation(w)$ (467), this becomes:

$$A!w \,\&\, \forall F(wF \rightarrow Propositional(F)) \,\&\, w\Sigma Fx_1 \ldots x_n$$

Expanding the definition of $Propositional(F)$ (275), this becomes:

$$A!w \,\&\, \forall F(wF \rightarrow \exists p(p = [\lambda y\, p])) \,\&\, w\Sigma Fx_1 \ldots x_n$$

Expanding the definition of $w\Sigma Fx_1 \ldots x_n$ (295), this becomes:

$$A!w \,\&\, \forall F(wF \rightarrow \exists p(p = [\lambda y\, p])) \,\&\, w\downarrow \,\&\, w[\lambda z\, Fx_1 \ldots x_n]$$

Expanding the definition of $w\downarrow$ (20.1), this becomes:

$$A!w \,\&\, \forall F(wF \rightarrow \exists p(p = [\lambda y\, p])) \,\&\, \exists F(Fw) \,\&\, w[\lambda z\, Fx_1 \ldots x_n]$$

Thus, $x_1,\ldots,x_n$ do not occur in encoding position of this last formula, and so by the Encoding Formula Convention (17.3), these variables do not occur in encoding position in the formula $w \models Fx_1 \ldots x_n$.

- Since the definite descriptions of our language are *rigid designators*, we've called expressions of the form $\imath x\varphi$ *rigid descriptions*.

However, it should be clear that *every* significant term of our language is a rigid designator, as this notion is generally understood. Semantically speaking, once an interpretation of the language and an assignment to the variables is specified, the denotation function is unary - the denotations of the terms are *not* relativized to possible worlds. So a significant term denotes the same entity in every modal context.

In addition to this notion of rigidity, we have introduced both metatheoretical notions (about the system) and theoretical notions (within the system) of rigidity. These notions, reviewed below, have a connection to the notion of *modal collapse*, which was discussed in Chapter 9, Section 9.9.3. A modally collapsed formula φ is one that can take any of a number of well-recognizable forms, such as $\Box(\varphi \to \varphi)$, $\Diamond\varphi \equiv \Box\varphi$, etc. Normally, if one can derive such formulas by modally strict means, then the modal collapse doesn't hinge on any contingency. See the theorems in Section 9.9.3.

In light of this, the metatheoretical notions of rigidity that we've introduced are:

- In (260), we defined φ is a *rigid condition on* α if and only if $\vdash_\Box \forall\alpha(\varphi \to \Box\varphi)$, i.e., if and only if it is a modally strict theorem that for any entity α, if φ, then necessarily φ.

- In (340), we defined φ is a *rigid restriction condition on* α if and only if φ is a restriction condition on α (336) and also a rigid condition on α (260). Recall that a restriction condition on α was defined in (336) as having the following characteristics:

  - φ has one free variable α,
  - $\vdash_\Box \exists\alpha\varphi(\alpha)$, i.e., it is provable, by modally strict means, that φ is non-empty, and
  - $\vdash_\Box \varphi^\tau_\alpha \to \tau\downarrow$, i.e., it is provable, by modally strict means, that φ has existential import.[298]

- In (340) we introduced *rigid restricted variables* to range over entities that satisfied a rigid restriction condition.

In contrast to these metatheoretical notions, we've defined a theoretical notion of rigidity in (571.1):

[298] By contrast, a *weak* restriction is condition was defined in (336) as any formula ψ such that (a) ψ has a single free variable α, (b) it is provable that ψ is *non-empty*, but not necessarily by modally strict means, and (c) it is provable that ψ has *existential import*, but not necessarily by modally strict means.

- We say that an n-ary relation R is *rigid* just in case $\Box\forall x_1 \ldots \forall x_n(Rx_1 \ldots x_n \rightarrow \Box Rx_1 \ldots x_n)$.

And in (571.2), we said that:

- A relation F^n *rigidifies* a relation G^n just in case F^n is rigid and is materially equivalent to G.

It may be that this summary will help one to cleanly separate the various notions of rigidity.

(**573**) **Theorem:** Every Relation Has a Rigidification. We first prove some lemmas. (.1) $x_1, \ldots x_n$ exemplify F-at-w if and only if $F^n x_1 \ldots x_n$ is true at w:

$$(.1)\ \ F^n_w x_1 \ldots x_n \equiv w \models F^n x_1 \ldots x_n \qquad (n \geq 0)$$

It then easily follows that (.2) G^n_w is rigid:

$$(.2)\ \ Rigid(G^n_w) \qquad (n \geq 0)$$

But an even more interesting theorem is implied, namely, that some relation rigidifies G^n:[299]

$$(.3)\ \ \exists F^n(Rigidifies(F^n, G^n)) \qquad (n \geq 0)$$

Note that (.3) can't be inferred immediately from (.2)! One can't use G^n_w where w is a fixed, but arbitrary, possible world to rigidify G. If you suppose G^n_w with $w = \boldsymbol{w}_\alpha$ is the witness to (.3), then the proof that $G^n_{\boldsymbol{w}_\alpha}$ is materially equivalent to G would fail to be modally strict. Nor would one be able to produce a relation that rigidifies G within an arbitrary modal context. Interestingly, though, the proof of (.3) does make use of G^n_w, but its use has to be set up properly.

By RN and GEN, it follows from (.3) that $\forall G \Box \exists F^n Rigidifies(F^n, G^n)$. We'll show in the discussion of (574.3) below that (.3) is a non-schematic, second-order version of an axiom formulated by Gallin (1975, 77).

Note that (.3) isn't derivable from the fact that every relation G^n has an actualized version $[\lambda x_1 \ldots x_n \mathcal{A}G^n x_1 \ldots x_n]$ that rigidifies G^n. We can see why if we consider the fact, noted above, that the present theorem holds necessarily for any property G. Now if we think semantically for the moment in terms of primitive possible worlds, this last fact means that for any possible world w and any relation G^n, there is a relation F^n whose exemplification extension at *every* possible world is the exemplification extension of G^n at w. By contrast, $[\lambda x_1 \ldots x_n \mathcal{A}G^n x_1 \ldots x_n]$ is a relation whose exemplification extension at *every* possible world is the exemplification extension of G^n at the actual world.

[299]Credit goes to Daniel Kirchner for noting (personal communication: 5 March 2020) that the following is a consequence of the Kirchner Theorem, reported in (271). In previous versions of object theory, this was taken as an axiom.

(**574**) **Theorems:** Equivalent Notions of Rigid Relations. Where $n \geq 0$, it is provable that (.1) the claim, (a) necessarily, for all objects $x_1, \ldots, x_n$, if $x_1, \ldots, x_n$ exemplify F^n, then necessarily $x_1, \ldots, x_n$ exemplify F^n, is equivalent to the claim, (b) for all objects $x_1, \ldots, x_n$, if possibly $x_1, \ldots, x_n$ exemplify F^n then necessarily $x_1, \ldots, x_n$ exemplify F^n; and (.2) the claim, (a) necessarily, for all objects $x_1, \ldots, x_n$, if $x_1, \ldots, x_n$ exemplify F^n, then necessarily $x_1, \ldots, x_n$ exemplify F^n, is equivalent to the claim, (b) for all objects $x_1, \ldots, x_n$, either necessarily $x_1, \ldots, x_n$ exemplify F^n or necessarily $x_1, \ldots, x_n$ fail to exemplify F^n:

(.1) $\Box\forall x_1 \ldots \forall x_n(Fx_1 \ldots x_n \rightarrow \Box Fx_1 \ldots x_n) \equiv \forall x_1 \ldots \forall x_n(\Diamond Fx_1 \ldots x_n \rightarrow \Box Fx_1 \ldots x_n)$

(.2) $\Box\forall x_1 \ldots \forall x_n(Fx_1 \ldots x_n \rightarrow \Box Fx_1 \ldots x_n) \equiv \forall x_1 \ldots \forall x_n(\Box Fx_1 \ldots x_n \vee \Box\neg Fx_1 \ldots x_n)$

Thus, it immediately follows from (.2), given definition (571.1), that (.3) F^n is rigid if and only if for all objects $x_1, \ldots, x_n$, either necessarily $x_1, \ldots, x_n$ exemplify F^n or necessarily $x_1, \ldots, x_n$ fail to exemplify F^n:

(.3) $Rigid(F^n) \equiv \forall x_1 \ldots \forall x_n(\Box Fx_1 \ldots x_n \vee \Box\neg Fx_1 \ldots x_n)$

(.3) is the definition of a rigid relation given in Montague 1974 (132) and in Gallin 1975 (77). Thus, theorem (573.3) captures, in second order form, the intuition underlying the typed axiom *Extensional Comprehension* formulated by Gallin (1975, 77).[300] Cf. Cocchiarella 1988 (54), ehere we find a similar, but non-modal, statement of what he calls the *principle of rigidity*.

(**575**) **Remark:** An Important Distinction. We are now in a position to consider the difference, if any, between the extension of a world-relativized property G_w and the world-relativized extension of G, and between the truth-value of a world-relativized proposition p and the world-relativized truth-value of p. Let's first consider the difference between extensions of world-relativized properties and world-relativized extensions of properties.

[300] Though we've captured the intuition in second-order form, a version of Gallin's typed axiom is derivable in typed object theory; see Chapter 15 of the present monograph for the presentation of typed object theory. For now, however, we can establish the connection with Gallin's principle as follows.

By applying RN and then GEN to (573.3), we obtain:

$$\forall G^n \Box \exists F^n(Rigidifies(F^n, G^n))$$

So by definition (571.2), this is equivalent to:

(ϑ) $\forall G^n \Box \exists F^n(Rigid(F^n) \,\&\, \forall x_1 \ldots \forall x_n(F^n x_1 \ldots x_n \equiv G^n x_1 \ldots x_n))$

Now in Gallin (1975, 77), the statement of (typed) *extensional comprehension* is:

$$\Box\exists \mathsf{f}_\sigma[\mathsf{Rn}(\mathsf{f}) \wedge \forall \mathsf{x}^0 \ldots \forall \mathsf{x}^{n-1}[\mathsf{f}\mathsf{x}^0 \ldots \mathsf{x}^{n-1} \leftrightarrow \mathsf{A}]],$$

where σ is the type of an n-ary relation among objects of arbitrary types, all the other expressions are appropriately typed, and f_σ doesn't occur free in A. Though (ϑ) is not a schema, it is a version of Gallin's principle that holds for n-ary relations among individuals. Note, finally, that a second-order schematic version of Gallix axiom, conditionalized on $[\lambda x_1 \ldots x_n\, \varphi]\downarrow$, can be derived using (569.2).

Since we've now established that G-at-w (G_w) exists (for any property G), it follows that its extension, ϵG_w, exists. And by (564.1) and (565), we know that the extension-at-w of G ($\epsilon_w G$) exists. So by applying definitions and theorems, we know:

$$\epsilon G_w = \imath x(A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv G_w z)))$$

$$\epsilon_w G = \imath x(A!x \,\&\, \forall F(xF \equiv w \models \forall z(Fz \equiv Gz)))$$

Intuitively, ϵG_w and $\epsilon_w G$ are distinct. If a property F is in fact materially equivalent to G-at-w, but not materially equivalent, at w, to G, then ϵG_w encodes F while $\epsilon_w G$ does not. And vice versa, if F is, at w, is materially equivalent to G, but is not in fact materially equivalent to G-at-w, then $\epsilon_w G$ encodes F while ϵG_w does not.

Exercise: Prove $\exists G \exists w(\epsilon G_w \neq \epsilon_w G)$. [Hint: Use (221.1).]

Now consider truth-values.

Since we've now established that p-at-w (p_w) exists (for any p), it follows that its extension, $\circ p_w$, exists. And by (556.2) and (557), we showed that the truth-value-at-w of p ($\circ_w p$) exists. So by applying definitions and theorems, we know:

$$\circ p_w = \imath s \forall q(s \models q \equiv (q \equiv p_w))$$

$$\circ_w p = \imath s \forall q(s \models q \equiv w \models (q \equiv p))$$

Intuitively, $\circ p_w$ and $\circ_w p$ are distinct. If a proposition q is in fact materially equivalent to p-at-w, but not materially equivalent, at w, to p, then $\circ p_w$ encodes q while $\circ_w p$ does not. And vice versa, if q is, at w, materially equivalent to p, but is not in fact materially equivalent to p-at-w, then $\circ_w p$ encodes q while $\circ p_w$ does not.

Exercise: Prove $\exists p \exists w(\circ p_w \neq \circ_w p)$.

12.4 Impossible Worlds

(**576**) **Remark:** On Impossible Worlds. From the 1960s through the 1990s, we start to find, in the literature, discussions of 'non-normal worlds' (Kripke 1965, Cresswell 1967, Rantala 1982, Priest 1992, and Priest & Sylvan 1992), 'non-classical worlds' (Cresswell 1972b), 'non-standard worlds' (Rescher & Brandom 1980, Paśniczek 1994) and 'impossible worlds' (Morgan 1973, Hintikka 1975, Routley 1980, Yagisawa 1988, Mares 1997, and Restall 1997). For a good overview of the recent literature on impossible worlds, see Berto 2013.[301]

[301] See also Nolan 2013, Krakauer 2013, and Jago 2013, which also focus on impossible worlds. However, these recent works don't consider the theory of impossible worlds developed in Zalta 1997a, which we further refine and develop here.

Though a variety of reasons have been given for postulating such impossible worlds, not all of those reasons are, from an object-theoretic perspective, convincing. For example, impossible worlds are often invoked to solve problems that arise when philosophers *represent* propositions as functions from possible worlds to truth-values (or as sets of possible worlds). Such representations, as is well known, *identify* propositions that are necessarily equivalent; if propositions p and q are just functions from worlds to truth-values, then they can't be distinguished when they have the same truth-value at every possible world. As a result, if one represents a belief as a relation between a person and a proposition so-conceived, then if x believes p, and p is necessarily equivalent to q, then x believes q. This just follows by the substitution of identicals and the identity of p and q. Such a result flies in the face of the data.[302] To solve this problem, it has been suggested that we can distinguish necessarily equivalent propositions if we consider their truth-values at impossible worlds. In effect, the suggestion is to represent propositions as functions from worlds generally, i.e., both possible and impossible worlds, to truth-values.

But, from the present perspective, we need not invoke impossible worlds to distinguish necessarily equivalent propositions. Object theory already has that capability; our theory of propositions doesn't collapse necessarily equivalent propositions. Thus, one can use the present theory of propositions to represent beliefs without incurring the result that one believes everything that is necessarily equivalent to what one believes.

Philosophers have also invoked impossible worlds to explain both impossibilities in fiction specifically and thoughts about impossible objects generally. But as the discussion in Section 12.6 will show, we can analyze such fictions and thoughts without invoking impossible worlds.

From an object-theoretic perspective, the best case for impossible worlds comes from:

(A) the analysis of counterfactual and subjunctive conditionals with impossible antecedents, and

(B) the study of paraconsistent logic.

As to (A), consider the following sentences, which are clearly true:

- If Frege's system had been consistent, he would have died a happier man.
- If there were a set of all non-self-membered sets, it would be a member of itself iff not a member of itself.

[302] Not only are there numerous examples of believing p without believing propositions necessarily equivalent to q, the problem of logical omniscience arises for this understanding of propositions. See Hintikka 1975 for a discussion of the problem and the suggestion that impossible worlds solve the problem.

The first example is a counterfactual conditional. On the standard analysis, such conditionals are true just in case the consequent is true at the closest possible world where the antecedent is true (Stalnaker 1968, Lewis 1973). But the antecedent of the example is true at no possible world, since one can derive a contradiction from Frege's axioms and thereby demonstrate his system's inconsistency. This inconsistency is not contingent; there is no possible world where the particular axioms of Frege's system are consistent. Hence, the antecedent of the above counterfactual conditional describes an impossibility and so the standard analysis implies that the sentence is false, thereby failing to preserve its truth-value.[303] It has been suggested that this problem might be solved if we amend the analysis so that counterfactual conditionals are considered true just in case the consequent is true at the closest world (possible or impossible) where the antecedent is true.

The second case is a subjunctive conditional, and there are lots of similar examples, such as, "if four were prime, it would be divisible only by itself and one". Again, on the standard analysis of the truth conditions of such subjunctive conditionals, on which the consequent holds in the closest possible world where the antecedent holds, the sentence turns out to be false, contrary to intuition. But though there is no possible world where something is a set of all non-self-membered sets or where four is prime, it is claimed that there are impossible worlds where these propositions are true. Of course, these claims are typically just assumed rather than proved to be true. By amending the analysis so that the truth conditions become "the consequent holds at the closest world (possible or otherwise) where the antecedent holds", we seem to get correct truth conditions for the subjunctive conditional, assuming that the notion of an impossible world is sufficiently clear.

Concerning (B), the suggestion that paraconsistent logic governs impossible worlds is persuasive. Paraconsistent logic weakens classical logic so that an arbitrary proposition can't be derived from the contradiction $p \,\&\, \neg p$; thus, the principle *ex contradictione (sequitur) quodlibet* fails for such a logic. On the analysis developed below, we can see why one might think there is a connection between paraconsistent logic and impossible worlds: from the fact that $p \,\&\, \neg p$ is true at an impossible world, it doesn't follow that every proposition is true at that world.

Of course, once one accepts that impossible worlds help us to understand

[303]Of course, one could try to 'paraphrase away' the description "Frege's system", by interpreting "Frege's system" as the *non-rigid* description "the system Frege developed", so that the sentence in question implies: if Frege had developed a consistent system, he would have died a happier man. Here, the antecedent is possibly true; there are possible worlds where Frege developed a consistent system. But that is not what the original sentence implies. The paraphrase is not an accurate one; it just misrepresents what the original sentence means. If this is not convincing, just change the example to: if the system Frege in fact developed had been consistent, he would have died a happier man.

the data presented by (A) and (B), the question arises, what are impossible worlds exactly? Too frequently, the answer is given by switching to model theory and modeling impossible worlds as sets of propositions. But it is important not to mistake the entities being modeled for the entities doing the modeling. A world, whether possible or impossible, is not a set of propositions. The propositions in a set don't characterize that set, whereas the propositions true in a world, possible or impossible, characterize the world.

In what follows, however, we develop a series of definitions and theorems that show impossible worlds to be abstract objects, indeed situations, characterized by the propositions true at them. Moreover, the most important principles governing impossible worlds are derived as theorems, such as the fundamental theorem that for every way a world couldn't possibly be, there is a non-trivial impossible world that is that way; this is proved in (585) below. I don't know of any other *theory* of impossible worlds that yields similar consequences.[304]

The reader may wish to consult Zalta 1997a for a more detailed motivation of impossible worlds than the one just presented. The work below revises, corrects, and enhances the theorems and proofs first developed there.

(**577**) **Definition and Theorems:** Impossible Worlds. In what follows, we continue to use the variables $s, s', s'', \ldots$ as rigid restricted variables ranging over situations (467). Recalling the definition of *Maximal*(s) (520) and the definition of *Possible*(s) (502), we may say that a situation s is an *impossible world* just in case s is maximal and not possible:

(.1) $\mathit{ImpossibleWorld}(s) \equiv_{df} \mathit{Maximal}(s) \,\&\, \neg \mathit{Possible}(s)$

i.e., if we eliminate the restricted variable:

$$\mathit{ImpossibleWorld}(x) \equiv_{df} \mathit{Situation}(x) \,\&\, \mathit{Maximal}(x) \,\&\, \neg \mathit{Possible}(x)$$

In this definition, maximality is a necessary condition. We take it that only those situations that are maximal are correctly considered to be worlds. Situations that are maximal and possible are (provably) possible worlds (522), while situations that are maximal and impossible are impossible worlds by definition. A situation that isn't maximal has no legitimate claim to being called a 'world', at least not in the technical sense philosophers attach to this notion in the attempt to understand modality.

Recall that we defined *TrivialSituation*(x) as a situation in which every proposition is true (487.2). After it was shown that $\imath x \mathit{TrivialSituation}(x)\downarrow$ (488.4) and $\boldsymbol{s_V}$ was defined as $\imath x \mathit{TrivialSituation}(x)$ (489.2), we gave a modally strict

[304] One way to question the arguments in Nolan 2013, Krakauer 2013, and Jago 2013, is to enquire whether their preferred theories of impossible worlds can yield the basic principles about impossible worlds as theorems, in the manner of Zalta 1997a and below.

proof that *TrivialSituation*($\boldsymbol{s_V}$) (490.4). It follows from these facts that (.2) $\boldsymbol{s_V}$ is an impossible world:

(.2) *ImpossibleWorld*($\boldsymbol{s_V}$)

(488.4) and (.2) imply that:

(.3) $\exists x$*ImpossibleWorld*(x)

and it is also straightfoward to show, for any individual term κ:

(.4) *ImpossibleWorld*(κ) $\rightarrow \kappa\downarrow$

Thus, *ImpossibleWorld*(x) is a restriction condition, as defined in (336). But before we introduce restricted variables for impossible worlds, we first show that it is a rigid restriction condition.

(**578**) **Theorems:** Modal Collapse of *ImpossibleWorld*(x) and Rigid Restricted Variables. It is a modally strict theorem that if x is an impossible world it is necessarily an impossible world:

(.1) *ImpossibleWorld*(x) $\rightarrow \Box$*ImpossibleWorld*(x)

The following are immediate consequences:

(.2) $\Box$(*ImpossibleWorld*(x) $\rightarrow \Box$*ImpossibleWorld*(x))

(.3) $\forall x$(*ImpossibleWorld*(x) $\rightarrow \Box$*ImpossibleWorld*(x))

So by (.2), *ImpossibleWorld*(x) is modally collapsed and, by (.3) and metadefinition (340), is a rigid restriction condition. Given the latter, we may introduce $i, i', i'', \ldots$ as (singly) restricted variables ranging over them. Moreover, we may take these new variables as doubly-restricted, i.e., as variables that range over situations s. Thus, at our convenience, we formulate theorems and reason with $i, i', \ldots$ as either singly restricted or doubly restricted, as discussed in Remark (514) for the case of the variables $w, w', w'', \ldots$. As a result, notions defined on situations s may be applied to impossible worlds i without having to be redefined.

(**579**) **Definition:** Truth at an Impossible World. An important example is the notion of *truth in* a situation, i.e., $s \models p$. By interpreting i as a doubly restricted variable, we may henceforth suppose that $i \models p$ is defined. We shall read $i \models p$ as 'p is true *at* i'. Thus, *truth at* an impossible world is simply a special case of the notion *truth in* a situation.

Note that if x is an impossible world, then it is a situation and so by definition (470) and theorem (88.8.i), $x \models p \equiv x\Sigma p$. Moreover, if x is an impossible world, then the fact that it is a situation implies x is abstract. So by similar reasoning, $x\Sigma p \equiv x[\lambda y\, p]$. Hence, when x is a situation, $x \models p \equiv x[\lambda y\, p]$. So for impossible worlds, $i \models p \equiv i[\lambda y\, p]$. Thus, a proposition p is true at i if and only if the property $[\lambda y\, p]$ *characterizes* i by way of an encoding predication. So we have a genuine theory of impossible worlds, not an ersatz theory or a set-theoretic model of them.

(**580**) **Theorem:** Identity of Impossible Worlds. Since impossible worlds are a species of situation, it follows that $i = i'$ iff all and only the propositions true at i are true at i':

$$i = i' \equiv \forall p(i \models p \equiv i' \models p)$$

(**581**) **Theorem:** The False is an Impossible World That Isn't Trivial. Recall the definition of The False ($\bot$) in (302.2). It follows by modally strict reasoning that The False is a non-trivial impossible world:

$$ImpossibleWorld(\bot)\,\&\,\neg TrivialSituation(\bot)$$

So there are non-trivial impossible worlds. Since $\bot$ is an impossible world at which every falsehood is true, we might take the liberty of referring to it as *the worst of all non-trivial impossible worlds*! Depending on your view of contradictions, the trivial situation, $\boldsymbol{s_V}$ (489.2), may be worse, since every proposition, as well as its negation, is true there. But then, $\boldsymbol{s_V}$ has a saving grace, for though every false proposition is true there, every true proposition is true there as well.

Exercise: In (559.2) we defined $\bot_w$ as The False at w. Develop a modally strict proof that, for any possible world w, $\bot_w$ is a non-trivial impossible world.

(**582**) **Theorem:** The False is an Impossible World That Is Not 1-Modally Closed. It is a modally strict theorem that $\bot$ is not is not (unary-) closed under necessary implication:[305]

$$\neg 1\text{-}ModallyClosed(\bot)$$

(**583**) **Definition:** The p-Extension of Situation s. We define *the p-extension of* situation s, written s^{+p}, as the situation that makes true all of the propositions that s makes true and also makes p true:

$$s^{+p} \ =_{df} \ \imath s'\forall q(s' \models q \equiv (s \models q \vee q = p))$$

Note that the expression s^{+p} is a binary functional term that involves two variables: the situation that s^{+p} denotes depends on the value of s and p. Since every 0-ary relation term Π^0 is significant, we may regard $\kappa^{+\Pi^0}$ as significant whenever κ is an individual term that is known, by hypothesis or by proof, to

[305] I'm indebted to Daniel West for suggesting that this theorem could be proved by modally strict means.

be such that $Situation(\kappa)$. For example, substituting $\perp$ for s and p_0 for p (where p_0 is $\forall x(E!x \rightarrow E!x)$) produces $\perp^{+p_0}$, which is provably a situation in which every false proposition is true and p_0 is true.

(**584**) **Lemmas:** The p-Extension of s is Strictly Canonical. Clearly, s^{+p} is a canonical situation, as this latter notion was reformulated in (486). Where φ is the formula $s \models q \vee q = p$, it follows, by modally strict means, that (.1) every proposition q such that φ is necessarily such that φ:

(.1) $\forall q((s \models q \vee q = p) \rightarrow \Box(s \models q \vee q = p))$

Hence s^{+p} is also a strictly canonical situation, as this notion was reformulated in (486), and so it is a modally strict consequence of (583) and (261.2) that (.2) the p-extension of s makes q true if and only if either s makes q true or q is identical to p:

(.2) $\forall q(s^{+p} \models q \equiv (s \models q \vee q = p))$

Thus, it immediately follows that: (.3) if a proposition is true in s it is true in the p-extension of s, and (.4) p is true in the p-extension of s:

(.3) $s \models q \rightarrow s^{+p} \models q$

(.4) $s^{+p} \models p$

(**585**) **Theorem:** Fundamental Theorem of Impossible Worlds. The most important fact about impossible worlds is that if it is not possible that p, then there is a non-trivial impossible world at which p is true:

$$\neg\Diamond p \rightarrow \exists i(\neg TrivialSituation(i) \,\&\, i \models p)$$

cf. Zalta 1997a, 647. If we borrow a turn of phrase from Lewis (1986, 2), we might read the above as: every way a world couldn't possibly be is a way some non-trivial impossible world is. Indeed, Nolan (1997, 542) suggests that a 'comprehension' principle governing impossible worlds should assert: for every proposition which cannot be true, there is an impossible world where that proposition is true. The present theorem clearly validates this principle, but not by stipulating it.

The proof of the above theorem in the Appendix corrects an error in the proof of the corresponding theorem in Zalta 1997a.[306]

[306] In Zalta 1997a (647–8), we correctly asserted, but incorrectly proved, the theorem that $\neg\Diamond p \rightarrow \exists s(ImpossibleWorld(s) \,\&\, s \neq s_u \,\&\, s \models p)$. In that theorem, s_u is the *universal* situation (i.e., the trivial situation we're now calling $\boldsymbol{s_V}$, in which every proposition is true). The proof developed in 1997a correctly established that $\neg\Diamond p$ implies that there is a situation s which is maximal, not possible, and in which p is true. But the proof that s is distinct from s_u contained an assumption that appeared correct but that isn't in fact provable, namely, that the impossible proposition p mentioned in the

The strategy of the proof in the Appendix is as follows: assume $\neg\Diamond p$; cite (531.1), i.e., that some possible world encodes all the truths; consider an arbitrary such possible world and then show that its p-extension is a non-trivial impossible world where p is true.

(**586**) **Theorems:** *Ex Contradictione Quodlibet* Fails for Impossible Worlds. Recall that we established, in (552.2), that the law *ex contradictione quodlibet* governs possible worlds. Formulated in the material mode, the law asserts: $w \models (p \,\&\, \neg p) \to w \models q$. But we can show that this law fails for impossible worlds, i.e., we can show: (.1) there are impossible worlds i and propositions p and q such that $(p \,\&\, \neg p)$ is true at i but q fails to be true at i:

(.1) $\exists i \exists p \exists q (i \models (p \,\&\, \neg p) \,\&\, \neg i \models q)$

Clearly, then, *ex contradictione quodlibet* fails for impossible worlds, as it does for situations generally.

Moreover, we can also reason from the fundamental theorem for impossible worlds (585) to show that a variant version of *ex contradictione quodlibet* fails for impossible worlds, i.e., that (.2) there are impossible worlds i and propositions p and q such that both p and $\neg p$ are true at i but q fails to be true at i:

(.2) $\exists i \exists p \exists q (i \models p \,\&\, i \models \neg p \,\&\, \neg i \models q)$

This variant of *ex contradictione quodlibet* thus fails for impossible worlds, as it does for situations generally. The point, however, is that both (.1) and (.2) identify *worlds* that validate paraconsistent logics in which *ex contradictione quodlibet* fails.

(**587**) **Theorem:** Disjunctive Syllogism Fails for Impossible Worlds. A version of disjunctive syllogism (86.4.b) governs possible worlds and propositions, for we established in (553) that: $(w \models (p \vee q) \,\&\, (w \models \neg p)) \to w \models q$. However, disjunctive syllogism fails for impossible worlds; there are impossible worlds i and propositions p and q such that $p \vee q$ and $\neg p$ are true at i but q fails to be true at i:

antecedent is distinct from the proposition $p \,\&\, \neg p$. The general claim that $q \neq (q \,\&\, \neg q)$ is certainly provable whenever q is a necessary, true, or even possible proposition. (For example, suppose $\Diamond q$, and assume for reductio that $q = (q \,\&\, \neg q)$. Then $\Diamond(q \,\&\, \neg q)$, which contradicts the fact that $\neg\Diamond(q \,\&\, \neg q)$.) However, when q is necessarily false, i.e., impossible, we can't prove the inequality $q \neq (q \,\&\, \neg q)$ from our axioms. Though our system requires that there be at least one impossible proposition (see (208.2)), it leaves open the question of whether there are multiple ones. Of course, it is consistent with the theory to assert that $q \neq (q \,\&\, \neg q)$ when q is necessarily false. If one does assert this, one can prove the existence of many new propositions. But when q is necessarily false, the identity $q = (q \,\&\, \neg q)$ is also consistent with the theory, and so one can't assume its negation. These facts were overlooked in the proof of the theorem in Zalta 1997a. By contrast, the proof of the present theorem (585) (i.e., in the Appendix) has been amended accordingly. In the corrected proof, the impossible world where p is true is shown to be non-trivial by identifying a *contingently false* proposition that fails to be true in it.

$$\exists i \exists p \exists q(i \models (p \vee q) \ \&\ i \models \neg p \ \&\ \neg i \models q)$$

Clearly, not only does disjunctive syllogism fail for impossible worlds but it also fails generally for situations. In any case, we have now shown that there are (impossible) worlds that validate logics in which disjunctive syllogism fails.

12.5 Moments of Time and World-States

(**588**) **Remark:** How Minimal Tense Logic Leaves Us Short of a Goal of Temporalization.[307] The definitions and theorems governing possible worlds that were developed in Section 12.2 can be adapted in a natural way to systematize *moments of time*, *world-states*, and other temporal abstractions. However, there are different ways of achieving this goal and they are not all equally satisfactory.

The main question is whether to start with the primitive operators of minimal tense logic, namely $\mathcal{H}$ ('it was always the case that') and $\mathcal{G}$ ('it will always be the case that'), or start with a primitive omnitemporal operator $\blacksquare$ ('it is always the case that'). Before we consider these options, note that in either case, it would be helpful to have an *intuitive* semantic picture by constructing a formal semantics for these operators. Both options require that we extend formal interpretations $\mathcal{I}$ of our language by adding a domain of times **T** that contains a distinguished element $\boldsymbol{t}_0$ (= the present moment). Moreover, if we start with the operators of minimal tense logic, we'll also need a binary relation, say < (= *occurs before*), on **T**. Then we would need to relativize the **ext**, **enc**, and **ex** functions to both world-time pairs, so that:

- **ext** would assign to each n-ary relation ($n \geq 1$) an exemplification extension at each world-time pair,

- **enc** would assign to each n-ary relation ($n \geq 1$) an encoding extension at each world-time pair, and

- **ex** would assign each to each 0-ary relation a truth-value at each world-time pair.

Formally, we would define $\mathbf{ext}_{\boldsymbol{w},\boldsymbol{t}}(\boldsymbol{r}^n)$, $\mathbf{enc}_{\boldsymbol{w},\boldsymbol{t}}(\boldsymbol{r}^n)$, and $\mathbf{ex}_{\boldsymbol{w},\boldsymbol{t}}(\boldsymbol{r}^0)$ as follows:

- $\mathbf{ext}_{\boldsymbol{w},\boldsymbol{t}} : \mathbf{R}_n \times \mathbf{W} \times \mathbf{T} \to \wp(\mathbf{D}^n)$ $\quad (n \geq 1)$

- $\mathbf{enc}_{\boldsymbol{w},\boldsymbol{t}} : \mathbf{R}_n \times \mathbf{W} \times \mathbf{T} \to \wp(\mathbf{D}^n)$ $\quad (n \geq 1)$

- $\mathbf{ex}_{\boldsymbol{w},\boldsymbol{t}} : \mathbf{R}_0 \times \mathbf{W} \times \mathbf{T} \to \{\boldsymbol{T}, \boldsymbol{F}\}$

[307] I'm indebted to Daniel West for his work on the question of how to temporalize object theory? His efforts forced me to think more deeply about the consequences of using minimal tense logic.

As one might expect, the definition of φ is *true with respect to* interpretation $\mathcal{I}$ and assignment f *at* world-time pair $\langle \boldsymbol{w}, \boldsymbol{t} \rangle$, written $\boldsymbol{w}, \boldsymbol{t} \models_{\mathcal{I},f} \varphi$, would have to be adjusted in a straightforward way, to account for the fact that we are evaluating truth with respect to world-time pairs.[308] The recursive clauses for the complex formulas $\neg\varphi$, $\varphi \to \psi$, and $\forall\alpha\varphi$ would be adapted similarly.

Now if we were to decide that the minimal tense logic operators should be primitive, we would use $<$ to state their truth conditions, relative to an interpretation $\mathcal{I}$ and assignment function f, as follows:

- $\boldsymbol{w}, \boldsymbol{t} \models_{\mathcal{I},f} \mathcal{H}\varphi$ if and only if $\forall \boldsymbol{t}'(\boldsymbol{t}' < \boldsymbol{t} \to \boldsymbol{w}, \boldsymbol{t}' \models_{\mathcal{I},f} \varphi)$
- $\boldsymbol{w}, \boldsymbol{t} \models_{\mathcal{I},f} \mathcal{G}\varphi$ if and only if $\forall \boldsymbol{t}'(\boldsymbol{t} < \boldsymbol{t}' \to \boldsymbol{w}, \boldsymbol{t}' \models_{\mathcal{I},f} \varphi)$

If we instead to decide that the omnitemporal operator $\blacksquare$ should be the sole primitive, we wouldn't need $<$ to state truth conditions, since these could be stated as:

- $\boldsymbol{w}, \boldsymbol{t} \models_{\mathcal{I},f} \blacksquare\varphi$ if and only if $\forall \boldsymbol{t}'(\boldsymbol{w}, \boldsymbol{t}' \models_{\mathcal{I},f} \varphi)$

Now there are a number of considerations that bear upon the choice of primitive operators. But before we examine those, there is another choice point that needs attention first, namely, how to assign truth conditions to the classical necessity operator $\Box$ when we move to a tensed version of object theory.

There are two options for understanding the notion of *metaphysically necessary* ($\Box$) semantically. One is to understand it narrowly as a universal quantifier just over semantically primitive possible worlds (Kaplan 1977 [1989], 545; Fine 1977), while the other is to understand it more broadly as a universal quantifier over semantically primitive world-time pairs (Montague 1973). That is, we have the following two options:

T8a. If φ is a formula of the form $\Box\psi$, then $\boldsymbol{w}, \boldsymbol{t} \models_{\mathcal{I},f} \varphi$ if and only if $\forall \boldsymbol{w}'(\boldsymbol{w}', \boldsymbol{t} \models_{\mathcal{I},f} \psi)$

T8b. If φ is a formula of the form $\Box\psi$, then $\boldsymbol{w}, \boldsymbol{t} \models_{\mathcal{I},f} \varphi$ if and only if $\forall \boldsymbol{w}' \forall \boldsymbol{t}'(\boldsymbol{w}', \boldsymbol{t}' \models_{\mathcal{I},f} \psi)$

Clearly on T8b, $\Box\varphi$ semantically implies that $\blacksquare\varphi$, but T8a has no such implication.

---

[308] Specifically:

T1. If φ is a formula of the form $\Pi\kappa_1 \ldots \kappa_n$ $(n \geq 1)$, then $\boldsymbol{w}, \boldsymbol{t} \models_{\mathcal{I},f} \varphi$ if and only if $\exists \boldsymbol{r}^n \exists \boldsymbol{o}_1 \ldots \exists \boldsymbol{o}_n(\boldsymbol{r} = \boldsymbol{d}_{\mathcal{I},f}(\Pi) \,\&\, \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f}(\kappa_1) \,\&\, \ldots \,\&\, \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f}(\kappa_n) \,\&\, \langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n \rangle \in \mathbf{ext}_{\boldsymbol{w},\boldsymbol{t}}(\boldsymbol{r}))$

T2. If φ is a formula of the form $\kappa_1 \ldots \kappa_n \Pi^n$ $(n \geq 1)$, then $\boldsymbol{w}, \boldsymbol{t} \models_{\mathcal{I},f} \varphi$ if and only if $\exists \boldsymbol{o}_1 \ldots \exists \boldsymbol{o}_n \exists \boldsymbol{r}^n(\boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f}(\kappa_1) \,\&\, \ldots \,\&\, \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f}(\kappa_n) \,\&\, \boldsymbol{r} = \boldsymbol{d}_{\mathcal{I},f}(\Pi) \,\&\, \langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n \rangle \in \mathbf{enc}_{\boldsymbol{w},\boldsymbol{t}}(\boldsymbol{r}))$

T3. If φ is a 0-ary relation constant or 0-ary relation variable Π, then $\boldsymbol{w}, \boldsymbol{t} \models_{\mathcal{I},f} \varphi$ if and only if $\mathbf{ex}_{\boldsymbol{w},\boldsymbol{t}}(\boldsymbol{d}_{\mathcal{I},f}(\Pi)) = \boldsymbol{T}$

Much can be said about these two options, but to immediately see one practical difference between them, reconsider the definition of a possible world as a situation x that *might* be such that all and only true propositions are true in x (512). Consider the definiens in (512):

$$Situation(x) \,\&\, \Diamond\forall p(x \models p \equiv p)$$

On the narrow interpretation of necessity ($\Box$) in T8a, as a universal quantifier just over possible worlds, possibility ($\Diamond$) becomes an existential quantifier just over possible worlds. So the formula displayed above is appropriate as a definiens for the definiendum *PossibleWorld*(x). However, if we take the wider interpretation of necessity as in T8b, as a universal quantifier over all world-time pairs, then possibility becomes an existential quantifier over such pairs. In that case, the definiens displayed above is more appropriate as a definiens for the definiendum *PossibleWorldState*(x).

But this is only one issue to keep in mind as one approaches the decision of how narrowly or widely to interpret $\Box$. Though in previous works (1987, 1988) I've assumed we should use something like T8a, the work in Dorr & Goodman 2019 and Daniel West (m.s.) develops a number of good reasons for preferring T8b. The main consideration discussed in Dorr & Goodman 2019 is how to preserve the intuition that *if a propositions is metaphysically necessary, it is always true* (i.e., $\forall p(\Box p \rightarrow \blacksquare p)$), and they give persuasive arguments that using T8b to interpret $\Box$ the better choice philosophically. West (m.s.) prefers T8b to T8a after specifically examining the relevant considerations that arise for object theory.

I plan to leave the decision about how to understand metaphysical necessity for others to resolve, since my goal here is to simply identify the important issues that arise when modifying the present theory to account for tense. In what follows, we only need to know that object theory would be extended by the theorem $xF \rightarrow \blacksquare xF$, either because we've opted for T8a and taken it as as axiom or because we've opted for T8b and it can be derived from $xF \rightarrow \Box xF$. (Either way, the more general claim, $x_1 \ldots x_n F \rightarrow \blacksquare x_1 \ldots x_n F$ will be derivable.) Moreover, before we return to the main question, i.e., whether to take the operators of minimal tense logic ($\mathcal{H}$, $\mathcal{G}$) or the omnitemporal operator ($\blacksquare$) as basic, it should also be mentioned that however we answer this question, it is of interest and importance to additionally introduce the operator *it is now the case that* ($\mathcal{N}$), which is to function tense-theoretically the way the actuality operator ($\mathcal{A}$) functions modally. That is, the truth conditions for the *now* operator are that $\mathcal{N}\varphi$ is true just in case φ is true at the distinguished time $\boldsymbol{t}_0$. Formally:

$$\boldsymbol{w}, \boldsymbol{t} \models_{\mathcal{I},f} \mathcal{N}\varphi \text{ if and only if } \boldsymbol{w}, \boldsymbol{t}_0 \models_{\mathcal{I},f} \varphi$$

With this in mind, let's first consider an issue that would arise if we were to take the operators of minimal tense logic as primitive. In discussing this issue, we'll need to define the duals of the operators $\mathcal{H}$ and $\mathcal{G}$, namely $\mathcal{P}$ ('it was once the case that') and $\mathcal{F}$ ('it will at some point be the case that'), as follows:

$$\mathcal{P}\varphi \equiv_{df} \neg\mathcal{H}\neg\varphi$$

$$\mathcal{F}\varphi \equiv_{df} \neg\mathcal{G}\neg\varphi$$

Let's suppose for the moment that we can successfully axiomatize these notions using the usual axioms for minimal tense logic for $\mathcal{H}$ and $\mathcal{G}$, which are:

$$\mathcal{H}(\varphi \to \psi) \to (\mathcal{H}\varphi \to \mathcal{H}\psi)$$

$$\mathcal{G}(\varphi \to \psi) \to (\mathcal{G}\varphi \to \mathcal{G}\psi)$$

$$\varphi \to \mathcal{H}\mathcal{F}\varphi$$

$$\varphi \to \mathcal{G}\mathcal{P}\varphi$$

Let us further assume that axioms for the now operator $\mathcal{N}$ are analogues of the axioms governing the actuality operator $\mathcal{A}$. But let's put aside, at least for the moment, the subtleties that *any* temporalization of object theory will have to address, such as (a) what axioms govern the interaction of the minimal tense operators and $\mathcal{N}$ with the other operators of the system, and (b) how to accommodate the *temporally fragile* axiom for the $\mathcal{N}$ operator ($\mathcal{N}\varphi \to \varphi$), and (c) accommodate temporally strict and non-strict derivations and proofs. These issues can be put aside because we are going to develop a distinctive problem for temporalizing object theory by taking $\mathcal{H}$ and $\mathcal{G}$ as primitive.

Before we get to this, however, let us briefly digress to say something more specific about what one might accomplish with the minimal tense operators. Note that the increased expressivity of tensed object theory allows one to formally represent the truth conditions and logical consequences of a wider variety of sentences in natural language, namely, those involving tenses. But, since we've thus far focused on the proof-theoretic power of object theory, the main goal is define philosophically interesting abstractions and prove philosophically interesting truths that are known *a priori*, as follows. In minimal tense logic, one may define the *omnitemporality* operator $\blacksquare$ (we'll henceforth read $\blacksquare\varphi$ as '*always* φ' instead of the more cumbersome '*it is always the case that* φ'). Moreover, its dual operator $\blacklozenge$ is definable (henceforth we'll read $\blacklozenge\varphi$ as '*sometime* φ' instead of the more cumbersome '*it is sometime the case that* φ'). The omnitemporal operator can be defined as:

$$\blacksquare\varphi \equiv_{df} \mathcal{H}\varphi \,\&\, \varphi \,\&\, \mathcal{G}\varphi$$

and it dual $\blacklozenge\varphi$ can be defined using either of the following equivalent definitions:

$$\blacklozenge\varphi \equiv_{df} \neg\blacksquare\neg\varphi$$

$$\blacklozenge\varphi \ \equiv_{df} \ \mathcal{P}\varphi \vee \varphi \vee \mathcal{F}\varphi$$

We may then use $\blacklozenge$ to say that a situation s is a *moment of time* just in case sometime, all and only true propositions are true in s:

$$\mathit{MomentOfTime}(s) \ \equiv_{df} \ \blacklozenge\forall p(s \models p \equiv p)$$

And if we use 't' as a restricted variable ranging over moments of time and use $t \models p$ (which is ultimately defined as $t[\lambda y\, p]$) to assert that p is true in t, we can say that a moment of time t is present if and only if every proposition true in t is true:

$$\mathit{Present}(t) \ \equiv_{df} \ \forall p((t \models p) \to p)$$

Clearly, these definitions are analogous to the definitions of *PossibleWorld*(s) and *Actual*(s).

If the changes needed to implement the above ideas could be successfully made to the system, then a good reason to adopt such principles and definitions would be to derive interesting and important philosophical theorems that govern times and world-states. For example, one might want to derive the claims: (.1) every moment of time is maximal; (.2) every moment of time t is such that sometime, t is present; (.3) there is a unique present moment of time; (.4) sometime p if and only if p is true at some moment of time; and (.5) always p if and only if p is true at every moment of time:

(.1) $\forall t \mathit{Maximal}(t)$

(.2) $\forall t \blacklozenge \mathit{Present}(t)$

(.3) $\exists! t \mathit{Present}(t)$

(.4) $\blacklozenge p \equiv \exists t(t \models p)$

(.5) $\blacksquare p \equiv \forall t(t \models p)$

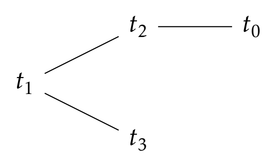
There are, in addition, many other notions that one might want to derive, for example, past and future moments of time:

$$\mathit{PastMoment}(s) \ \equiv_{df} \ \mathcal{P}\forall p(s \models p \equiv p)$$

$$\mathit{FutureMoment}(s) \ \equiv_{df} \ \mathcal{F}\forall p(s \models p \equiv p)$$

All of this looks like a promising way to extend object theory into the temporal domain.

However, there is a distinctive philosophical problem (a mismatch of expectations) that arises with this approach. To get a proper understanding of this problem, it helps to reconsider the intuitive semantics introduced earlier, in which we helped ourselves to some set theory, primitive moments of time, and primitive world-time pairs. Note the following fact about minimal tense logic: the axioms of minimal tense logic don't require that the semantic relation $<$ be an *ordering* of the elements of the semantic domain **T** of primitive times. That is, $<$ need not be irreflexive, asymmetric, and transitive. Indeed, $<$ need not even be a connected (linear) relation – there may distinct times $\boldsymbol{t}$ and $\boldsymbol{t}'$ (i.e., $\boldsymbol{t} \neq \boldsymbol{t}'$) in the domain **T** such that neither $\boldsymbol{t} < \boldsymbol{t}'$ nor $\boldsymbol{t}' < \boldsymbol{t}$.[309] Using these facts about minimal tense logic, we can state the problem (though it will take some reasoning to make it clear): in any interpretation in which the relation $<$ is not transitive and linear, there are primitive times in the semantic domain **T** that aren't represented by an abstract situation s that satisfies the defined condition *MomentOfTime*(s). Consider any model (i.e., interpretation in which the axioms of minimal tense logic are true) based on a temporal frame in which the domain **T** consisted of 4 pairwise-distinct times, $\boldsymbol{t}_0$ – $\boldsymbol{t}_3$, the only facts about which are:



$\boldsymbol{t}_1 < \boldsymbol{t}_2$

$\boldsymbol{t}_2 < \boldsymbol{t}_0$

$\boldsymbol{t}_1 < \boldsymbol{t}_3$

Note that since $\boldsymbol{t}_0$ is the present moment, there are no 'future' times in this temporal frame (i.e., there are no times $\boldsymbol{t}$ such that $\boldsymbol{t}_0 < \boldsymbol{t}$) and there is only one time, namely $\boldsymbol{t}_2$, in the past of the present moment $\boldsymbol{t}_0$ (since neither $\boldsymbol{t}_1 < \boldsymbol{t}_0$ nor $\boldsymbol{t}_3 < \boldsymbol{t}_0$). Moreover, (a) the transitivity of $<$ fails because $\boldsymbol{t}_1 < \boldsymbol{t}_2$, $\boldsymbol{t}_2 < \boldsymbol{t}_0$, and $\neg(\boldsymbol{t}_1 < \boldsymbol{t}_0)$, and (b) the connectivity (linearity) of $<$ fails in two different ways: $\neg(\boldsymbol{t}_1 < \boldsymbol{t}_0 \vee \boldsymbol{t}_0 < \boldsymbol{t}_1)$ and $\neg(\boldsymbol{t}_2 < \boldsymbol{t}_3 \vee \boldsymbol{t}_3 < \boldsymbol{t}_2)$.

Now if this temporal frame is used as a basis for a model of object theory, then we cannot regard the entities in the domain of times **T** as abstract objects. Both $\boldsymbol{t}_1$ and $\boldsymbol{t}_3$ are counterexamples. Both are primitive times in the domain **T**, but neither can be regarded as an abstract object that satisfies the definition

[309] Moreover:

- since $\mathcal{H}\varphi \to \varphi$ and $\mathcal{G}\varphi \to \varphi$ are not axiomatic, the binary relation $<$ need not be reflexive
- since $\mathcal{H}\varphi \to \mathcal{H}\mathcal{H}\varphi$ and $\mathcal{G}\varphi \to \mathcal{G}\mathcal{G}\varphi$ are not axiomatic, the binary relation $<$ need not be transitive
- since $\mathcal{H}\mathcal{H} \to \mathcal{H}\varphi$ and $\mathcal{G}\mathcal{G}\varphi \to \mathcal{G}\varphi$ are not axiomatic, $<$ need not be dense
- and so on, for the other principles that might place conditions on the relation $<$.

of *MomentOfTime*(s) (i.e., $\blacklozenge\forall p((s \models p) \equiv p)$). To see why, evaluate the definition at $\boldsymbol{t}_0$. To satisfy the definition of *moment of time* at $\boldsymbol{t}_0$, a situation s has to be such that "sometime, s makes true all and only the propositions p that are true (at that time)". Here, the operator "sometime", when evaluated at $\boldsymbol{t}_0$, is witnessed only by those primitive times $\boldsymbol{t}$ in the domain such that either $\boldsymbol{t} < \boldsymbol{t}_0$ or $\boldsymbol{t} = \boldsymbol{t}_0$ or $\boldsymbol{t}_0 < \boldsymbol{t}$. So $\boldsymbol{t}_1$ and $\boldsymbol{t}_3$ are times in the domain that won't be a witness to the definition. Thus, we can't regard them as abstract objects that satisfy the definition.

Of course, one could *add*, to minimal tense logic, the axioms that would guarantee the transitivity and linearity (connectedness) of the relation $<$ on the domain of times:

Transitivity: $\mathcal{H}\varphi \to \mathcal{H}\mathcal{H}\varphi$
$\mathcal{G}\varphi \to \mathcal{G}\mathcal{G}\varphi$

Linearity (Connectedness): $(\mathcal{P}\mathcal{F}\varphi \vee \mathcal{F}\mathcal{P}\varphi) \to \blacklozenge\varphi$

But if one shares the intuition that a moment of time should be identifiable as an abstraction independent of any particular temporal structure, then the above method of temporalizing object theory doesn't preserve this intuition. In the next Remark, we consider an alternative method, one that lets us define moments of time as abstractions without requiring any conditions on the structure of time. We'll see below that there is a way to extend object theory into the temporal domain without being forced into making choices about the structure of time.[310]

(**589**) **Remark:** A Hybrid Tense Logic. In light of the foregoing Remark, the simplest way of extending object theory into the temporal domain is to do the following:

(A) Add a primitive omnitemporal operator $\blacksquare$, with its defined dual $\blacklozenge$, and axiomatize them exactly as the $\Box$ and $\Diamond$ are axiomatized in classical quantified S5 modal logic. In other words, we might need, in the first instance at least, all the closures (including omnitemporal closures, i.e., the result of prefacing any instance by any string of $\blacksquare$ operators) of the K, T, and 5 axioms:

- $\blacksquare(\varphi \to \psi) \to (\blacksquare\varphi \to \blacksquare\psi)$

[310] Earlier versions of this monograph (prior to 2022) didn't flag this problem; I discovered it while reading Daniel West's (unpublished) attempts to work out some of the details of the approach. Again, I'm indebted to Daniel for his efforts, since they forced me to think more deeply about the consequences of using minimal tense logic.

Also, I've benefited from reading Prior 1968, Kamp 1971, S. K. Thomason 1972, Fine 1977, Burgess 1984, R. Thomason 1984, Wölfl 1999, Meyer 2009, Blackburn & Jorgensen m.s., and especially Goranko &: Rumberg 2022.

- $\blacksquare\varphi \rightarrow \varphi$
- $\blacklozenge\varphi \rightarrow \blacksquare\blacklozenge\varphi$

This allows us to derive BF and CBF in the usual way, so that $\forall\alpha\blacksquare\varphi \equiv \blacksquare\forall\alpha\varphi$ and $\blacklozenge\exists\alpha\varphi \equiv \exists\alpha\blacklozenge\varphi$ become theorems. A temporal counterpart for axiom (45.4) is not needed – it is not an *a priori* matter whether *sometime, there is a concrete object that is not now concrete*. However, an additional axiom is needed, for the commutativity of the $\Box$ and $\blacksquare$, namely, the closures of the following:

- $\Box\blacksquare\varphi \equiv \blacksquare\Box\varphi$

The system described thus far could be interpreted over a semantic domain of times that includes a distinguished present moment $\boldsymbol{t}_0$ but is otherwise unstructured (i.e., doesn't satisfy the conditions of any $<$ relation) much as the $\Box$ operator, under T8a, would be interpreted over an unstructured domain of worlds with a distinguished world $\boldsymbol{w}_0$.

(B) Note that we do *not* want definite descriptions to have denotations that can vary with the temporal context. We want them to be rigid designators, both modally and temporally, so that the substitution of identicals will work in any context. Thus, in order to properly axiomatize definite descriptions, we would have a use for the 'Now' operator $\mathcal{N}$ in the language and an axiom analogous to (43)$\star$ must be added, namely $\mathcal{N}\varphi \rightarrow \varphi$. This becomes a *temporally fragile* axiom and so one must refrain from asserting the omnitemporal closures of $\mathcal{N}\varphi \rightarrow \varphi$. Moreover, we take all the *presentization closures* (i.e., the result of prefacing any instance by any string of $\mathcal{N}$ operators) of all the other, non-temporally fragile axioms discussed thus far.

(C) In light of (B), the axiom governing descriptions (47) needs to be revised so that $\imath x\varphi$ denotes an object uniquely such that $\mathcal{A}\mathcal{N}\varphi$. It should read:

$x = \imath x\varphi \;\equiv\; \forall z(\mathcal{A}\mathcal{N}\varphi^z_x \equiv z = x)$, provided z is substitutable for x in φ and doesn't occur free in φ

Thus $\imath x\varphi$ would have a new significance; the description $\imath x\varphi$ would be read as 'the x (in fact, at present) such that φ'. A semantics for this operator is straightforward.[311] The axiom for descriptions only needs the

[311] Formally, if given an interpretation $\mathcal{I}$ and assignment f, we may specify the denotation$_{\mathcal{I},f}$ of $\imath x\varphi$, written $\boldsymbol{d}_{\mathcal{I},f}(\imath x\varphi)$, as follows, where $\boldsymbol{o}$ and $\boldsymbol{o}'$ are arbitrary objects in the domain of individuals and $f[x/\boldsymbol{o}]$ is the assignment just like f except that it assigns the object $\boldsymbol{o}$ to the variable x:

$$\boldsymbol{d}_{\mathcal{I},f}(\imath x\varphi) = \begin{cases} \boldsymbol{o}, \text{ if } \boldsymbol{w}_0, \boldsymbol{t}_0 \models_{\mathcal{I},f[x/\boldsymbol{o}]} \varphi \;\&\; \forall\boldsymbol{o}'(\boldsymbol{w}_0, \boldsymbol{t}_0 \models_{\mathcal{I},f[x/\boldsymbol{o}']} \varphi \rightarrow \boldsymbol{o}' = \boldsymbol{o}) \\ \text{undefined, otherwise} \end{cases}$$

placement of the $\mathcal{N}$ operator immediately after the $\mathcal{A}$ operator. It should also be an axiom or theorem that the actuality and now operators commute, i.e., that $\mathcal{A}\mathcal{N}\varphi \equiv \mathcal{N}\mathcal{A}\varphi$. This change may make it impossible to regard temporal object theory as a simple extension of modal object theory.

(D) The presence of the operator $\mathcal{N}$ requires axioms governing its interactions with the other operators of the system, and the natural way to do this is to recast the necessary axioms of $\mathcal{A}$ in (44) as omnitemporal axioms for $\mathcal{N}$ and, depending on whether one adopts T8a or T8b, recast the axioms for the interaction of $\mathcal{A}$ and $\Box$ in (46) as axioms for the interaction of $\mathcal{N}$ and $\blacksquare$ operator. [312] It is important to remember that the truth of $\mathcal{N}\varphi$ is always evaluated semantically at $\boldsymbol{t}_0$ and that one should *not* suppose that this distinguished (semantically primitive) time *was* a future time and *will be* a past time. It is inappropriate to use primitive temporal locutions like 'was' and 'will be' to assert intuitions about the semantically primitive time $\boldsymbol{t}_0$ – the whole point of the semantics is to cache out primitive temporal locutions like 'was' and 'will be' in terms of *atemporal* entities in the semantics. And if we've set up object theory correctly, this also applies to the unique present moment definable in object theory.[313]

[312]For the former, we need:

$\mathcal{N}\neg\varphi \equiv \neg\mathcal{N}\varphi$

$\mathcal{N}(\varphi \to \psi) \equiv (\mathcal{N}\varphi \to \mathcal{N}\psi)$

$\mathcal{N}\forall\alpha\varphi \equiv \forall\alpha\mathcal{N}\varphi$

$\mathcal{N}\varphi \equiv \mathcal{N}\mathcal{N}\varphi$

For the latter, we need:

$\mathcal{N}\varphi \to \blacksquare\mathcal{N}\varphi$

$\blacksquare\varphi \equiv \mathcal{N}\blacksquare\varphi$

[313]Of course, we may *define* temporal notions in object theory and apply them to times just as we have defined modal notions in object theory and applied them to possible worlds. For example, in (502), we defined *Possible*(s) as $\Diamond$*Actual*(s), where *Actual*(s) is defined as $\forall p(s \models p \to p)$ (492). Then it follows from (517) by GEN that every possible world is possible $\forall w Possible(w)$, and so every possible world is possibly actual $\forall w\Diamond Actual(w)$. That captures the intuition some philosophers and logicians have when they say, about semantically primitive possible worlds, that every such world might have been actual. But whereas such claims are problematic if these philosophers are using primitive modal notions to assert facts about primitive possible worlds, our claim that every world is possibly actual uses the *defined* notion of actuality and so doesn't make the mistake of using primitive modal notions to attribute modal properties to semantically primitive possible worlds.

Similarly, we may define *Historical*(s) as $\blacklozenge$*Present*(s), where *Present*(s) is defined as $\forall p(s \models p \to p)$. Then it follows that every moment of time is historical, i.e., $\forall t Historical(t)$, and so every moment of time present sometime $\forall t\blacklozenge Present(t)$. This, too, avoids the error of attributing primitive temporal properties to semantically primitive times.

Once this system is in place, one could then define, as described in the previous remark:

$$MomentOfTime(s) \equiv_{df} \blacklozenge\forall p(s \models p \equiv p)$$

This definition would be satisfied by all and only the elements of the primitive semantic domain of times. Moreover, truth at a time then becomes defined in the usual way: where t is an abstract object satisfying the definition just given $t \models p$ would become defined as $t\Sigma p$, i.e., as $t[\lambda y\, p]$. One can then define a *present moment* as any moment t such that $\forall p(t \models p \rightarrow p)$ and then prove that there is a unique present moment, as well as the usual counterparts of the theorems of world theory, such as the fundamental theorems: (a) *sometime* p if and only if there is a time at which p is true, i.e., $\blacklozenge p \equiv \exists t(t \models p)$, and (b) *always* p if and only if p is true at every time, i.e., $\blacklozenge p \equiv \forall t(t \models p)$.

Note that by defining moments of time, truth at a time, and introducing restricted variables to range over them, we have a hybrid logic and can make use of that to assert whichever the principles that we might want to adopt about the structure of time. Indeed, we may assert these principles as modally fragile axioms, so that we allow for possible worlds where time has a different structure. In particular, one can now extend the language of object theory by introducing a new object-theoretic relation $<$ on times t (i.e., on abstract objects that satisfy the definition displayed above) and axiomatize that relation however we like. One has any number of axioms to choose from, such as the following, in which $\leq$ is defined in the usual way (see Goranko and Rumberg 2022, Section 2.1):

- irreflexivity: $\forall t \neg(t < t)$
- transitivity: $\forall t\forall t'\forall t''(t < t' \,\&\, t' < t'' \rightarrow t < t'')$
- asymmetry: $\forall t\forall t' \neg(t < t' \,\&\, t' < t)$
- anti-symmetry: $\forall t\forall t'(t < t' \,\&\, t' < t \rightarrow t = t')$
- linearity (trichotomy, connectedness): $\forall t\forall t'(t < t' \vee t' < t \vee t = t')$
- density: $\forall t\forall t'(t < t' \rightarrow \exists t''(t < t'' \,\&\, t'' < t'))$
- forward-linearity: $\forall t\forall t'\forall t''(t'' < t \,\&\, t'' < t' \rightarrow (t = t' \vee t < t' \vee t' < t))$
- backward-linearity: $\forall t\forall t'\forall t''(t < t'' \,\&\, t' < t'' \rightarrow (t = t' \vee t < t' \vee t' < t))$
- beginning: $\exists t \neg\exists t'(t' < t)$
- end: $\exists t \neg\exists t'(t < t')$
- no beginning: $\forall t\exists t'(t' < t)$

- no end (unboundedness): $\forall t \exists t'(t < t')$
- forward-discrete:
  $\forall t_1 \forall t_2(t_1 < t_2 \rightarrow \exists t_3(t_1 < t_3 \,\&\, t_3 \leq t_2 \,\&\, \neg\exists t_4(t_1 < t_4 \,\&\, t_4 < t_3)))$
- backwards-discrete:
  $\forall t_1 \forall t_2(t_2 < t_1 \rightarrow \exists t_3(t_3 < t_1 \,\&\, t_2 \leq t_3 \,\&\, \neg\exists t_4(t_3 < t_4 \,\&\, t_4 < t_1)))$

Clearly, by asserting one or more of the above as modally fragile axioms, we can preserve such intuitions as: time is linear (dense, etc.) but might not have been!

Moreover, this method allows one to reconstruct the operators of minimal tense theory by definition. For example, where t_0 has been introduced as a name of the unique present moment, we may define:

$$\mathcal{H}\varphi \equiv_{df} \forall t(t < t_0 \rightarrow t \models \varphi)$$

Though I'm sure that I haven't anticipated all of the problems that may arise and need to be solved, I think this is the most promising course of action for temporalizing object theory. As far as I can see, it can achieve all the goals of temporalization while minimizing the costs and maximizing flexibility as to which principles about the nature of time should be adopted.

12.6 Stories and Fictional Individuals

We now explore the suggestion that *stories*, like *possible worlds* and *impossible worlds*, are a species of *situations* and then analyze the notions of *character* and *fictional character* in terms of stories. For the purposes of this section only, we assume that our language has been extended with individual constants representing proper names, historical names, names of stories, and names of fictional characters, and with relation names representing the names of ordinary properties and relations of the kind that appear in classical fiction. In this section we omit discussion of the data involving *fictional properties* (e.g., *being a unicorn*, *being a hobbit*) and *fictional relations* (e.g., *absolute simultaneity*). A discussion and analysis of these entities is reserved for Chapter 15, where we analyze fictional properties and fictional relations as higher-order abstracta, namely, abstract properties and abstract relations.

12.6.1 Data and Methodology

(**590**) **Remark:** Data To Be Explained. In what follows, we shall assume that there are at least some truths and valid inferences pertaining to fiction expressible in natural language. These 'data' fall into four types, where we use σ as a metalinguistic variable that ranges over names and descriptions of stories from natural language:

(A) pre-theoretic *truths* about stories and characters that would become falsehoods if prefaced by a locution of the form 'In the story' or 'According to the story';[314]

(B) true claims about the world that are taken to be true when authors use them in the context of a story;

(C) pre-theoretic truths that have the form 'In the story $\sigma, \ldots$' or 'According to the story $\sigma, \ldots$'; and

(D) pre-theoretic judgments about what logically follows from the truths in (A), (B), and (C).

The important point here is not that, for each type of data, there is some *fixed*, easily identifiable group of such truths or judgments, but only that there are truths and judgments of each kind. There will no doubt be disagreement as to which sentences constitute data of type (A) – (C), and disagreement about what exactly follows from these sentences. But our analysis won't hinge on there being agreement about this data, but only agreement that there are some data of the kinds just outlined. Our analysis will then be as precise as it can be *given* the disagreement about the data. In other words, our analysis will be precise but *open-ended*, in the sense that the specific conclusions that can be drawn from the analysis depend upon what data is agreed upon and subjected to the analysis. So there can be agreement about the analysis without there being agreement about what data is subject to the analysis.

Examples of (A) are:

(.1) (a) *The Iliad* is a story.
(b) *Crime and Punishment* is a story.
(c) *The Brother Karamazov* is a story.
(d) *A Study in Scarlet* is a story.

(.2) (a) Homer authored *The Iliad*.
(b) Dostoyevsky authored *Crime and Punishment*.
(c) Conan Doyle authored *A Study in Scarlet*.

(.3) Homer authored *The Iliad* before Dostoyevsky authored *Crime and Punishment*.

(.4) (a) Porphyry is an original character of *Crime and Punishment*.
(b) Sherlock Holmes is an original character of *A Study in Scarlet*.

[314] For the purposes of this work, I shall not distinguish 'In the story' and 'According to the story'. However, see Semeijn 2021, Ch. 7, for a way to draw a distinction between the two.

(.5) (a) Porphyry is a fictional character.
(b) Sherlock Holmes is a fictional character.

(.6) (a) Porphyry is a fictional detective.
(b) Holmes is a fictional detective.

(.7) (a) London is a character of *A Study in Scarlet*.
(b) London isn't an original character of *A Study in Scarlet*.
(c) London isn't a fictional character.
(d) The London of *A Study in Scarlet* is an original character of *A Study in Scarlet*.
(e) The London of *A Study in Scarlet* is a fictional character.
(f) The London of *A Study in Scarlet* is a fictional city.
(g) London is not identical to the London of *A Study in Scarlet*.

(.8) *Crime and Punishment* is a fiction, not a false story.

(.9) Sherlock Holmes is more famous than Porphyry.

Examples of (B) are:

(.10) (a) St. Petersburg is a city.
(b) Baker Street is in London.

Note that (a) remains true when prefaced by 'In *Crime and Punishmment*' but not when prefaced by 'In *The Iliad*', and that (b) remains true when prefaced by 'According to *A Study in Scarlet*' but not when prefaced by 'According to *The Brothers Karamazov*'.

Examples of (C) are:

(.11) (a) In *The Iliad*, Achilles fought Hector.
(b) According to *Crime and Punishment*, Raskolnikov kills a pawnbroker.
(c) According to *A Study in Scarlet*, Holmes is a detective.
(d) According to *A Study in Scarlet*, Holmes might have been a violinist.
(e) According to *Crime and Punishment*, Porphyry is a detective.
(f) In *A Study in Scarlet*, London is a city.

Examples of (D) are:

(.12) *The Iliad* is a story.
Therefore, there are truths according to *The Iliad*.

(.13) *Crime and Punishment* is a story.
Therefore, something authored *Crime and Punishment*.

(.14) In *Crime and Punishment*, Porphyry is a detective.
Porphyry is an original character of *Crime and Punishment*.
Therefore, Porphyry is a fictional detective.

(.15) London is a real city.
Therefore, London is not a fictional city.

(.16) London is a real city.
London is a character of *A Study in Scarlet*.
A Study in Scarlet is a fiction.
Therefore, some characters of a fiction are real.

(.17) The London of *A Study in Scarlet* is an original character of *A Study in Scarlet*.
In *A Study in Scarlet*, London is a city.
Therefore, the London of *A Study in Scarlet* is a fictional city.

(.18) London is a real city.
The London of *A Study in Scarlet* is a fictional city.
Therefore, London is not (identical to) the London of *A Study in Scarlet*.

(.19) Augustus Caesar worshipped Jupiter.
Jupiter is a fictional character.
Fictional characters aren't real.
Therefore, Augustus Caesar worshipped an object that isn't real.

(.20) Sherlock Holmes inspires some criminologists.
Sherlock Holmes is fictional.
Therefore, a fictional object inspires some criminologists.

(**591**) **Remark:** Methodology. Our methodology is to proceed in three stages:

(I) We introduce new theoretical notions (both primitive and defined) into object theory; these are notions needed for the analysis of the data.

(II) The data in (A), (B), and (C) of (590) are then analyzed by representing them in object theory using the notions introduced in (I). As part of the analysis, we index the names of the *original* characters of a story with the name of that story. The data, as analyzed in object theory, may then be used either as (modally fragile) premises and conclusions of arguments representable in object theory or as axioms of object theory. (For simplicity, we usually take them to be (modally fragile) premises or conclusions within a derivation, rather than as axioms.)

(III) We then show how the judgments of logical consequence in (D) of (590) are validated as *derivations* in object theory, i.e., we show, for each argument in (D), that the analysis of the conclusion is derivable from the analyses of the premises.

As to (I), we shall need three, primitive theoretical notions. We leave a full analysis of these notions for some other occasion since, for the most part, an exact axiomatization isn't essential to the analysis presented below. One of theses notions, temporal precedence, was discussed in the previous section and little more needs to be said about it. But a few minimal remarks (or assumptions) about the primitives are in order.

The three primitive theoretical notions required for the treatment of fiction are:

- a binary *authorship* relation ('A') among individuals
  'Axy' asserts: *x authors y*
- an $n+1$-ary *relevant entailment* operation ('$\Rightarrow_R$') on propositions $(n \geq 1)$
  '$(p_1,\ldots,p_n) \Rightarrow_R q$' asserts: *$p_1,\ldots,p_n$ relevantly imply q*
- a binary *temporal precedence* operation ('$<$') on propositions
  '$p < q$' asserts: *p temporally precedes q* (or *p before q*)

For the purposes of this chapter, we may regard formulas of the form Axy, $(p_1,\ldots,p_n) \Rightarrow_R q$, and $p < q$ as 0-place relation terms. While 'A' can clearly be added to our language as a distinguished binary relation term, we leave it to the reader to determine how best to extend the language and system of object theory with the binary operation symbol $<$ and $n+1$-ary operation symbol $\Rightarrow_R$.

The following, minimal understanding of our three new primitives is all that will be required:[315]

- *Authorship*. We rely on the intuition that an individual x may author a story y by way of a story-telling, where a story-telling can take place either by oral narration, by writing down sentence-tokens on paper, by using a keyboard on a computer to configure electonic bits that represent sentences, etc. Clearly, then, (.1) if y authors x, then y is a concrete object, and (.2) if something authors x, then there might not have been something that authors x. We therefore employ the following two minimal axioms for *authorship*:

  (.1) $Ayx \rightarrow E!y$

  (.2) $\exists yAyx \rightarrow \Diamond\neg\exists yAyx$

  (.2) captures the intuition that authored objects might not have been authored.

[315] The following serves to refine the discussion of these primitives found in Zalta 1983 (Chapter IV), Zalta 1988 (124–125), and 2000c (Section 4).

- *Relevant Entailment*. In what follows, we shall assume that (.3) the pretheoretic *story operator* is closed under relevant entailment, so that when stories are imported into object theory, the following principle holds, where $\underline{s}$ is a story, as defined in (592.1) below:

  (.3) $(\underline{s} \models p_1 \,\&\, \ldots \,\&\, \underline{s} \models p_n \,\&\, (p_1,\ldots,p_n \Rightarrow_R q)) \rightarrow \underline{s} \models q$

  In other words, any proposition relevantly implied by propositions true according to a story $\underline{s}$ are also true according to $\underline{s}$. Moreover, if we define $\varphi \Leftrightarrow_R \psi$ as $(\varphi \Rightarrow_R \psi) \,\&\, (\psi \Rightarrow_R \varphi)$, then we assert that (.4) β-Conversion holds for relevant entailment:

  (.4) $[\lambda x_1 \ldots x_n\, \varphi]\downarrow \rightarrow ([\lambda x_1 \ldots x_n\, \varphi]x_1 \ldots x_n \Leftrightarrow_R \varphi)$

  In other words, $[\lambda x_1 \ldots x_n\, \varphi]x_1 \ldots x_n$ and φ relevantly imply each other whenever $[\lambda x_1 \ldots x_n\, \varphi]$ exists. Note how (.3) and (.4) may combine. (.4) yields both $Rab \Rightarrow_R [\lambda x\, Rxb]a$ and $Rab \Rightarrow_R [\lambda x\, Rax]b$, since both $[\lambda x\, Rxb]$ and $[\lambda x\, Rax]$ exist. Then if given that "In $\underline{s}$, a bears R to b" ($\underline{s} \models Rab$), (.3) implies that "In $\underline{s}$, a exemplifies $[\lambda x\, Rxb]$" ($\underline{s} \models [\lambda x\, Rxb]a$) and "In $\underline{s}$, b exemplifies $[\lambda x\, Rax]$" ($\underline{s} \models [\lambda x\, Rax]b$).

  Intuitively, as we listen to, or read, a narrative, we don't infer all the logical consequences of the propositions signified by the sentences explicitly uttered or contained in the narrative, but instead infer only the consequences that are, in some sense, 'relevant' to forming a common sense understanding of the story being narrated (Parsons 1980, 175–82). Which consequences are relevant is highly contextual and depends on a number of factors, including the context in which the story is narrated by the author, the context in which the hearer or reader becames acquainted with the narrative, the cognitive state of the hearer or reader, how well the hearer or reader can pick up allusions, etc.[316] So, strictly speaking, the truth of fictional claims may depend on a context of utterance. Our analysis assumes that there are contexts in which fictional claims of the form "According to σ, φ" are true and thereby constitute data.

- *Temporal Precedence*. Intuitively, $p < q$ simply has to represent the meaning of the adverb 'before' in the following ordinary sentences:

  Caesar crossed the Rubicon before Washington crossed the Delaware.

[316] I take the present theory to be consistent with the fact that sometimes, e.g., when we are listening to an unreliable narrator, we might regard some propositions explicitly asserted in the story as ones that are not true in the story. And I also take the present theory to allow for the existence of truths in the story that arise in virtue of some formal aspects of the text or narration, as described in Kim 2022.

Homer authored *The Iliad* before Proust authored *Rememberance of Things Past*.

If we ignore the past tense of the verbs and use obvious abbreviations, then the former would be represented formally as $Ccr < Cwd$ and then second would be represented formally as $Ahi < Apr$.[317] Formally, the *temporal precedence* operator among propositions can be understood in terms of the notion of temporal precedence among moments of time discussed in the previous section. Where $t <_T t'$ is a relation on times that holds whenever time t precedes time t', then (.5) p *before* q just in case there are times t and t' such that t precedes t', p occurs at t and at no earlier time, and q occurs at t' and at no earlier time:

(.5) $p < q \equiv \exists t \exists t'(t <_T t' \;\&\; t \models p \;\&\; \neg\exists t''(t'' <_T t \;\&\; t'' \models p) \;\&\; t' \models q \;\&\; \neg\exists t''(t'' <_T t' \;\&\; t'' \models q))$

It is important to stress again that there may be no general agreement as to which propositions are true in a story or the exact logic of relevant entailment. Such lack of agreement won't matter, however, since we assume only that (i) there are at least *some* true judgments of the form "According to $\underline{s}$, . . . ", (ii) that there are some true judgments about what is relevantly entailed the truths of a story, and (iii) the judgments in (i) and (ii) may require us to invoke, and draw information from, the context in order to develop a precise analysis. If the reader denies the truth of the particular examples of story truths and entailments that we've chosen, they may simply substitute others, possibly by specifying a context. So when we say that the story operator is closed under relevant entailment, the reader may substitute any sentences of the form "In story $\underline{s}$, . . . " they judge to be true and close the resulting story operator under those relevant entailments they judge to hold.

This understanding of our three, new primitive notions should suffice for a basic analysis of fiction within object theory.

12.6.2 Principles For Analyzing Fiction

We therefore first implement (I) in (591) by introducing new definitions, axioms, and theorems of object theory.

(**592**) **Definition and Theorem:** Stories. In what follows, we take the variables x, y, z to be unrestricted. We now say that (.1) x is a *story* just in case x is a non-null situation that has an author:

[317] Though we've taken a liberty by ignoring the past tense in the two target sentences, note that we could represent the natural language sentences as $[\lambda\, Ccr] < [\lambda\, Cwd]$ and $[\lambda\, Ahi] < [\lambda\, Apr]$, respectively. These could be read without the past tense as: *that Caesar crosses the Rubicon* temporally precedes *that Washington crosses the Delaware*, and *that Homer authors* The Iliad temporally precedes *that Proust authors* Remembrance of Things Past.

(.1) $Story(x) \equiv_{df} Situation(x) \,\&\, \neg Null(x) \,\&\, \exists y Ayx$

Cf. Zalta 1983, p. 91, and 2000, p. 123. Note that since stories are situations, we might have expressed (.1) as $Story(s) \equiv_{df} \neg Null(s) \,\&\, \exists y Ays$.

Definition (.1) does *not* require stories to be *fictions*. For example, someone's testimony (whether true or false) in a court of law constitutes a story, as would a reporter's news article (again, whether true or false). Any authored narrative constitutes a story. Later we'll investigate what it takes for stories to be fictions and, to simplify matters, our attempt to answer this question will put aside the intentions of the author, so that we can give an answer based solely on a story's content; see (604).

The definition of a story immediately implies that stories have existential import, since the following is a (modally strict) theorem, for any individual term κ:

(.2) $Story(\kappa) \to \kappa\downarrow$

This follows from the fact that $Story(\kappa) \to A!\kappa$, by (.1) and the definition of a situation (467).

(**593**) ⋆**Axiom:** On the Existence of Stories. We shall, for the purposes of this section, take the claim that there exist stories to be a modally fragile axiom:

$$\exists x Story(x)$$

The reason for regarding the claim as modally fragile is worth discussing. By axiom (591.2), anything authored by something might not have been authored by something. But more importantly, given the additional reasonable assumption that it might not have been the case that there are objects y and x such that y authors x ($\Diamond\neg\exists y\exists x Ayx$), it follows that $\Diamond\neg\exists x Story(x)$.[318]

[318] Assume $\Diamond\neg\exists y\exists x Ayx$. Then by a modally strict theorem of quantification logic ($\exists y\exists x\varphi \equiv \exists x\exists y\varphi$) and a Rule of Substitution, we know:

(ϑ) $\Diamond\neg\exists x\exists y Ayx$

Note, independently, that it is a modally strict theorem that if $\neg\chi \to \neg(\varphi \,\&\, \psi \,\&\, \chi)$. As an instance, we have:

$$\neg\exists y Ayx \to \neg(Situation(x) \,\&\, \neg Null(x) \,\&\, \exists y Ayx)$$

By the Rule of Substitution for Defined Formulas, it follows that:

$$\neg\exists y Ayx \to \neg Story(x)$$

Since this holds for any x (i.e., x isn't free in any assumption), it follows by GEN that:

$$\forall x(\neg\exists y Ayx \to \neg Story(x))$$

So by quantification theory (39.3):

$\forall x\neg\exists y Ayx \to \forall x\neg Story(x)$, i.e.,
$\neg\exists x\exists y Ayx \to \neg\exists x Story(x)$

Since this is a modally strict theorem, it follows by RM$\Diamond$:

Consequently, the existence of stories isn't assertible *a priori* as a necessary truth. So if we want to add $\exists x Story(x)$ to object theory (not merely as an occasional assumption), then we must flag it as a modally fragile $\star$-axiom.

(**594**) **Remark:** Weak Restricted Variables for Stories. From (592.1), (593)$\star$, and (592.2), it follows that $Story(x)$ is a *weak restriction condition*, as defined in (336) – inspection of (592.1) shows that it has a single free variable; and (593)$\star$ establishes that it is provably (but not strictly provably) non-empty; and (592.2) establishes that it provably has existential import. So we can introduce $\underline{s}, \underline{s}', \underline{s}'', \ldots$ as *weak* restricted variables ranging over stories.

Since we established above that it is possible that there are no stories, then $Story(x)$ also fails to be a *rigid* restriction condition, as defined in (340). For the claims $\exists x Story(x)$ and $\Diamond\neg\exists x Story(x)$ imply $\neg\forall x(Story(x) \rightarrow \Box Story(x))$.[319] So there are multiple reasons why we can't regard our restricted variables for stories as *rigid* restricted variables. Indeed, from our axiom (591.2), that if y authors x then possibly y doesn't author x (i.e., $Ayx \rightarrow \Diamond\neg Ayx$), it follows that $\forall x(Story(x) \rightarrow \Diamond\neg Story(x))$.[320] Consequently, it is important to remember that if we assume $Story(\kappa)$ holds for some individual term κ, such an assumption has to be marked as modally fragile. The fact that $Story(\kappa)$ depends on a contingency, namely, that someone has engaged in behavior that constitutes authoring κ.

(**595**) **Remark:** Truth in a Story. The weak restricted variables that we introduced to range over stories can be interpreted as doubly-restricted, for reasons analogous to those discussed in (514). This means that notions defined on situations apply to stories without having to be redefined. So by interpreting $\underline{s}$ as doubly restricted, the claim that $\underline{s} \models p$ is defined. But whereas we read $s \models p$ as 'p is true in (situation) s', we shall read '$\underline{s} \models p$' as 'p is true according to (story) $\underline{s}$' or as 'in story $\underline{s}$, p'. This makes it clear which formulas of object theory serve to represent some of the data discussed earlier.

It is important to observe a critical difference between $s \models p$ and $\underline{s} \models p$, however. Whereas claims of the form $s \models p$ can be derived by modally strict means, claims of the form $\underline{s} \models p$ can not. The existence of situations can be proved by modally strict means, and whenever s is a situation, the propositions

$$\Diamond\neg\exists x\exists y Ayx \rightarrow \Diamond\neg\exists x Story(x))$$

But from this and (ϑ), it follows that $\Diamond\neg\exists x Story(x)$. $\bowtie$

[319] The proof is similar to the one in footnote 246.

[320] To see this, apply GEN to (591.2) so that we know $\forall x(\exists y Ayx \rightarrow \Diamond\neg\exists y Ayx)$. Since x isn't free in our assumption, it suffices by GEN to prove $Story(x) \rightarrow \neg\Box Story(x)$. So assume $Story(x)$. Then, by the definition of a story (592.1), it follows *a fortiori* that $\exists y Ayx$. So by (591.2), $\Diamond\neg\exists y Ayx$, i.e., $\neg\Box\exists y Ayx$. Now suppose $\Box Story(x)$, for reductio. Then, by the definition of a story (592.1) and the Rule of Substitution for Defined Formulas (160.2), it follows that:

$$\Box(Situation(x) \,\&\, \neg Null(x) \,\&\, \exists y Ayx)$$

A fortiori, $\Box\exists y Ayx$. Contradiction.

p true in s can be derived from the canonical description that identifies which object s is. But the existence of stories cannot be proved by modally strict means, and neither can claims of the form $\underline{s} \models p$. When we represent "σ is a story" and "According to $\sigma, \ldots$" as $Story(\sigma)$ and $\sigma \models \varphi$ (where φ represents the natural language claim filling in the ellipsis), then we have to regard the formal representations as modally fragile assumptions or modally fragile axioms (and marked with a $\star$). Consequently, any conclusion of the form $\Box\sigma \models p$ (where σ has been substituted for $\underline{s}$ in $\underline{s} \models p$) derived from the premise that $\sigma \models p$ will not be a modally strict conclusion; even though the derivation is modally strict, the conclusion $\Box\sigma \models p$ depends on a contingency.

(**596**) **Theorem:** Truth in a Story is a Rigid Condition on Properties and a Proviso. It is straightforward to establish that for any proposition p, if p is true in story $\underline{s}$, then necessarily, p is true in story $\underline{s}$, i.e.,

$$\forall p(\underline{s} \models p \rightarrow \Box\underline{s} \models p)$$

An interesting subtlety about this theorem arises once it is observed that true ordinary language statements of the form "According to $\sigma, \ldots$" are analytic. When an author introduces the proper name σ to title a story or introduces a proper name for new, original characters of the story, they are *changing the expressive power of the language*.[321] It then becomes important to regard claims of the form "According to $\sigma, \ldots$" as *analytic* truths about the language with the new expressive power. But, these are not ordinary analytic truths; since they are modally fragile, they constitute a class of analytic truths who representations, $\sigma \models p$, are provably necessary but the proof of which depends upon a contingency; $\Box\sigma \models p$ becomes derivable (by the above theorem) from $\sigma \models p$, but since the latter has to be flagged as modally fragile axiom or assumption, the necessary truth derived from it has to be similarly flagged as a non-modally strict theorem or conclusion.

In consequence, any attempt to instantiate the variable $\underline{s}$ in the above claim to a particular story, say σ, will result in a non-modally strict conclusion! For as we saw earlier, the assumption or axiom that $Story(\sigma)$ is modally fragile. If $Story(\sigma)$ is a modally fragile *assumption*, then although the inference to $Situation(\sigma)$ is a modally strict inference, the conclusion $Situation(\sigma)$ still depends on a non-modally strict assumption. And if we instead assert $Story(\sigma)$ as a modally fragile *axiom* instead of an assumption, then the theorem $Situation(\sigma)$ will not be a modally strict theorem. In the instance of the above theorem in which σ is substituted for $\underline{s}$, i.e., $\forall p(\varphi^{\sigma}_{\underline{s}} \rightarrow \Box\varphi^{\sigma}_{\underline{s}})$, can not be established by modally strict means.

[321] For example, when Conan Doyle wrote *A Study in Scarlet*, this and other facts (e.g., *A Study in Scarlet* appeared in book form in 1888, Sherlock Holmes is a fictional character, etc.), became expressible; they weren't expressible prior to the storytelling.

A fuller discussion of this point will be reserved for the discussion of (Leibnizian) concepts in Chapter 13; see especially Remark (695) in Section 13.3, where we discuss a class of necessary truths the proofs of which are not modally strict because they depend upon a contingency.

(**597**) **Theorem:** The Identity of Stories. From our definitions, it follows that a story $\underline{s}$ is the situation s' that makes true just the propositions p such that p is true in $\underline{s}$:

$$\underline{s} = \imath s' \forall p(s' \models p \equiv \underline{s} \models p)$$

Thus, stories can be generally identified as strictly canonical situations, but as noted in the discussion following the previous theorem, any attempt to instantiate this claim to particular story σ results in a non-modally strict instance, since the inference will depend on the modally fragile claim that $Story(\sigma)$. Indeed, once we represent natural language data of the form "According to σ, ..." as modally fragile assumptions of the form $\sigma \models p$, any conclusions we derive about which properties are encoded by σ will similarly be non-modally strict conclusions.

(**598**) **Definition:** Characters of Stories. For the purposes of this section, it suffices to say that *x is a character of* $\underline{s}$ just in case there is some property that x exemplifies according to $\underline{s}$:

$$CharacterOf(x, \underline{s}) \equiv_{df} x\downarrow \,\&\, \exists F(\underline{s} \models Fx)$$

Though this definition will need to be refined to handle a special case (see below), note that it allows real objects to be characters of stories. For suppose we're given the claims: (a) in *A Study in Scarlet* (σ_1), Baker Street exemplifies the property of being located in London, i.e., $\sigma_1 \models [\lambda x\, Lxl]b$, and (b) in *A Study in Scarlet*, London exemplifies the property of being the location of Baker Street, i.e., $\sigma_1 \models [\lambda x\, Lbx]l$. Then it follows from (a) that there is a property that Baker Street exemplifies in *A Study in Scarlet*, i.e., $\exists F(\sigma_1 \models Fb)$, and it follows from (b) that there is a property that London exemplifies in *A Study in Scarlet*, i.e., $\exists F(\sigma_1 \models Fl)$. Hence, by the above definition, it follows, respectively, that Baker Street is a character of *A Study in Scarlet*, i.e., $CharacterOf(b, \sigma_1)$, and that London is a character of *A Study in Scarlet*, i.e., $CharacterOf(l, \sigma_1)$.

Of course, if one treats occurrences of 'London' and 'Baker Street' in *A Study in Scarlet* as denoting original, fictional characters, namely 'The London of *A Study in Scarlet*' and 'The Baker Street of *A Study in Scarlet*', respectively, then we may not conclude anything about London and Baker Street; they are not characters of *A Study in Scarlet*, though of course they exemplify some properties that the fictional characters encode. We explore the details of this in the next subsection.

For most purposes and most cases of classical fiction, the above definition of *CharacterOf* suffices. But if one wants to forestall the problem of *named* characters that are distinct but indiscernible, the above definition has to be refined a bit. We leave this to a later chapter.[322]

(**599**) **Definition:** Originate In and Original Characters. (.1) An object x *originates in* a story $\underline{s}$ just in case x is a character of $\underline{s}$ that is abstract and that is not a character of any story authored before $\underline{s}$ was authored; moreover (.2) x is an *original character of* $\underline{s}$ just case x originates in $\underline{s}$:

(.1) $OriginatesIn(x,\underline{s}) \equiv_{df} CharacterOf(x,\underline{s}) \,\&\, A!x \,\&$
$\forall y \forall z \forall \underline{s}'((Az\underline{s}' < Ay\underline{s}) \rightarrow \neg CharacterOf(x,\underline{s}'))$

(.2) $OriginalCharacterOf(x,\underline{s}) \equiv_{df} OriginatesIn(x,\underline{s})$

(**600**) **Axiom:** Identity of Original Characters. We now stipulate that if an object x is an original character of $\underline{s}$, then x encodes exactly those properties that x exemplifies according to $\underline{s}$:

$$OriginalCharacterOf(x,\underline{s}) \rightarrow x = \imath y(A!y \,\&\, \forall F(yF \equiv \ \underline{s} \models Fy))$$

Intuitively, if Raskolnikov is an original character of *Crime and Punishment*, then Raskolnikov is the abstract object that encodes just the properties F such that Raskolnikov exemplifies F according to *Crime and Punishment*.

(**601**) **Definition:** Fictional Characters. We say that an object is a fictional character just in case it is an original character of some story:

$$Fictional(x) \ \equiv_{df} \ \exists \underline{s} OriginalCharacterOf(x,\underline{s})$$

For example, this definition lets us infer, from the premise that Raskonikov is an original character of *Crime and Punishment*, that Raskolnikov is a fictional character.

(**602**) **Definition:** Fictional Gs. We say that *x is a fictional G* just in case x is an original character of a story according to which x exemplifies G:

[322] The issue here is similar to the problem of named indiscernibles in mathematics, e.g., i and $-i$ in complex number theory. Nodelman & Zalta 2014 (52) forestall the issue by defining '*x is an element of* the structure T' (or 'x is an object of theory T') as: it is true in T that every y that is T-distinct from x is distinguishable from x by some property. This utilizes a special *abstract relation* $=_T$, for each theory T. Similarly, once we have developed typed object theory in Chapter 15, we may introduce, for each story $\underline{s}$, a special *abstract relation* of identity$_{\underline{s}}$ ($=_{\underline{s}}$). Then we can refine the above definition to read:

$$CharacterOf(x,\underline{s}) \ \equiv_{df} \ \underline{s} \models \forall y(y \neq_{\underline{s}} x \rightarrow \exists F(Fx \,\&\, \neg Fy))$$

In other words, x is a character of $\underline{s}$ just in case it is true in $\underline{s}$ that for every individual y that is distinct$_{\underline{s}}$ from x, x is exemplification-distinguishable from y by some property. A fuller discussion of this definition is left for Chapter 15.

$$Fictional\text{-}G(x) \equiv_{df} \exists \underline{s}(OriginalCharacterOf(x, \underline{s}) \,\&\, \underline{s} \models Gx)$$

(**603**) **Theorems:** Facts about Fictions. Theorems (222.3), (115.3), and (126.1) imply that (.1) if x is abstract, x couldn't be identical with any concrete object. It then follows that (.2) if x is fictional, then x is not identical with any possibly concrete object; and (.3) if x is a fictional-G, then x is not identical with any possibly concrete G:

(.1) $A!x \rightarrow \neg\exists y(\Diamond E!y \,\&\, y = x)$

(.2) $Fictional(x) \rightarrow \neg\exists y(\Diamond E!y \,\&\, y = x)$

(.3) $Fictional\text{-}G(x) \rightarrow \neg\exists y\big(\Diamond(E!y \,\&\, Gy) \,\&\, y = x\big)$

Cf. Zalta 2006a, 600-601. Now, if we read '$\Diamond E!y$' as y is a possible object, then as an applied instance of (.2), we know: if Sherlock Holmes is fictional then there is no possible object with which Holmes is identical, and as an applied instance of (.3), we know: if Sherlock Holmes is a fictional person, then there is no possible person with which Holmes is identical.

Consider how these applied instances of (.2) and (.3), together with the assumption that Holmes is fictional (or Holmes is a fictional person), allow us to derive, in precise terms, a view that Kripke developed in the following passage in his 1972 [1980]:

> I hold similar views about fictional proper names. ... Similarly, I hold the metaphysical view that, granted that there is no Sherlock Holmes, one cannot say of any possible person that he *would have been* Sherlock Holmes, had he existed. Several distinct possible people, and even actual ones such as Darwin or Jack the Ripper, might have performed the exploits of Holmes, but there is none of whom we can say that he would have *been* Holmes had he performed those exploits. For if so, which one?
>
> (1972 [1980, 157–158])

Of course, someone might argue that in the second sentence of the above quote, when Kripke assumes "there is no Sherlock Holmes", he meant only to suppose $\neg\exists x(x = h)$. Such an assumption implies that the name 'Sherlock Holmes' is empty. But that can't be what Kripke meant in the above passage, since he goes on to say that "one cannot say of any possible person that he would have been Holmes had he existed". One can't meaningfully make such a claim if 'Holmes' were an empty name, for how could a counterfactual identity statement ("he would have been Holmes") be true if one side of the identity statement has an empty name?

Moreover, there are two additional pieces of evidence for the suggestion that (.2) helps us to interpret this particular passage in Kripke. The first is the fact that in 1973 [2013], he takes at least some uses of fictional names to be non-empty. He says:

> A fictional character, then, is an abstract entity. It exists in virtue of the more concrete activities of telling stories, writing plays, writing novels, and so on, under criteria which I won't try to state precisely, but which should have their own obvious character. ... Of course, a fictional person isn't a person. (1973 [2013, 73–74])
>
> ...
>
> Let's take the statement 'Hamlet was a fictional character'. But applying the predicate ['was a fictional character'] on the level of reality—that is, so to speak, straight—one should say Hamlet *was* a fictional character. (1973 [2013, 74])
>
> ...
>
> Finally, at any rate, my view gives another sense in which it is true rather than false that Hamlet exists. At least, one should say 'There really is such a fictional character as Hamlet'. Such a fictional character really exists... (1973 [2013, 78])

I think Kripke is here committed to the claim that Holmes is fictional, and so (.2) captures his view that Holmes isn't identical with any possible (i.e., possibly concrete) object.

The second is the fact that Kripke makes an analogous claim/argument concerning fictional species such as *unicorns*, namely, that they can't be identified with any possible species. But further discussion of this second piece of evidence will have to be postponed until Chapter 15, Section 15.5.5, where we use a type-theoretic framework to show how to capture his analogous claim regarding fictional species; see (978.14) – (978.16) and the discussion there.

Finally, note Kripke's 'aside' in the above quotation from 1973 [2013], "Of course, a fictional person isn't a person". This, too, can be represented and derived in the present theory, given (a) the notion of a *concreteness-entailing property*, which we defined but didn't tag in (434), and the assumptions that (b) *being a person* ('P') is a concreteness-entailing property, and (c) Holmes is a fictional person. For if (c), then by (602), Holmes is an original character of some story. Hence, by (599.2) and (599.1), Holmes is abstract, and so necessarily non-concrete ($\Box\neg E!h$). But by (a), (b), and the definition of concreteness-entailing, it is easy to show $\Box\forall x(\neg E!x \rightarrow \neg Px)$. Hence, $\forall x\Box(\neg E!x \rightarrow \neg Px)$. Instantiating to Holmes: $\Box(\neg E!h \rightarrow \neg Ph)$. By the K axiom, $\Box\neg E!h \rightarrow \Box\neg Ph$. Since we've already established the antecedent, it follows that $\Box\neg Ph$, and so by the T schema, $\neg Ph$. That is, Holmes doesn't exemplify *being a person*, despite being a fictional person, which, I take it, is what Kripke is claiming. But now we have a derivation from general assumptions, instead of an assertion.

(**604**) **Definition:** Fictions vs. True and False Stories. We now distinguish between true stories, false stories, and fictions (i.e., fictional stories), based solely on their content. That is, in the discussion that follows, we shall purposefully ignore the intentions of the author. No doubt the intentions of an author often play a role in determining whether a story is true, false, or a fiction, and so the definitions proposed below should be seen as having a *ceteris paribus* ('all other things being equal') clause. But an author's intentions can't be the sole determiner of whether a story is fictional or not, since in the cases of many myths, legends, etc., the stories were told, and intended, as fact, yet we want to regard such myths and legends a fictions.

But our purpose now is to consider whether a story, for which it known only that it is an authored narrative, is a true story, a false story, or a fiction. Thus, the technical notions defined here may differ from the ordinary, everyday concept of fiction if the everyday concept requires that the intentions of an author play a role in determining whether a story is true, false, or a fiction. Indeed, the technical notions defined here may differ from the ordinary notions in other ways, since the everyday concept of a story, in some contexts, classifies false stories as fictions. Narratives that contains a false proposition (e.g., as in false testimony) are often labeled as 'fictions'. But as we shall use the term here, a fiction is not simply an authored (non-null) situation that encodes a false proposition; in addition, at least one of its characters must be a *fictional* character. By contrast, none of the characters in true and false stories may be fictional.

We therefore say: (.1) a story $\underline{s}$ is *true* just in case every proposition true in $\underline{s}$ is true and there are no original characters of $\underline{s}$; (.2) a story $\underline{s}$ is *false* just in case some proposition true in $\underline{s}$ fails to be true and there are no original characters of $\underline{s}$; and (.3) a story $\underline{s}$ is a *fiction* just in case something is an original character of $\underline{s}$:

(.1) $True(\underline{s}) \equiv_{df} \forall p(\underline{s} \models p \rightarrow p) \,\&\, \neg\exists x(OriginalCharacterOf(x, \underline{s}))$

(.2) $False(\underline{s}) \equiv_{df} \exists p(\underline{s} \models p \,\&\, \neg p) \,\&\, \neg\exists x(OriginalCharacterOf(x, \underline{s}))$

(.3) $Fiction(\underline{s}) \equiv_{df} \exists x(OriginalCharacterOf(x, \underline{s}))$

As noted above, these technical notions may differ from the ordinary, everyday concepts. That's because we're interested primarily in the ontological status of stories; we're not as much interested in the context in which the story is told (e.g., whether it is explicitly being put forward as a 'fiction') or the popular use of the term 'fiction' to label a narrative that contains a falsehood.

The present view is based on the idea that one can't simultaneously assert that a story s is both false and a fiction. Of course, many stories are *popularly* held to be both false and fictions, but here we are drawing a distinction between a false story, something that is truth-apt, and a fiction, something that isn't. Here is a rule of thumb. Suppose a name κ of a known concrete object y is used in the storytelling of $\underline{s}$ to name a character of $\underline{s}$. And suppose, further, that there is some claim p that is true in $\underline{s}$ but that isn't true *simpliciter*, and that $\underline{s}$ has no other obvious fictional characters. Then $\underline{s}$ is a false story if and only if κ, as used in $\underline{s}$, is a name of the concrete object y. Alternatively $\underline{s}$ is a fiction if κ, as used in $\underline{s}$, is a name of the original character 'the κ of $\underline{s}$'.

For example, consider a story, say $\underline{s}_1$, in which (a) the term 'Richard Nixon' is used in the storytelling of $\underline{s}_1$ to name the main character of $\underline{s}_1$, (b) $\underline{s}_1$ encodes a false proposition (say, that 'Nixon was born in Delaware'), and (c) none of the other characters of $\underline{s}_1$ is fictional. Then, we should take $\underline{s}_1$ to be a false story if and only if we take 'Richard Nixon', as used in the storytelling of $\underline{s}_1$, to denote the historical person who was the 37th President of the U.S. On the other hand, we should take $\underline{s}_1$ to be a fiction if and only if we take 'Richard Nixon', as used in the storytelling of $\underline{s}_1$, to denote an original (and hence, fictional) character describable as 'the Richard Nixon of $\underline{s}_1$'. But we shouldn't regard $\underline{s}_1$ as both a false story and a fiction.

12.6.3 Analysis of the Data

In this subsection, we implement step (II) of (591), by analyzing the data in (A), (B), and (C) of (590) in terms of the principles developed in the previous subsection.

(**605**) **Remark:** Analyses of the (A) Examples. In what follows, we analyze the (A) examples of (591). In cases where there are multiple examples with the same form, we give only one example and then produce the analysis. When we introduce a new constant to represent the name of an original character in story $\underline{s}$, we index the constant to the name of $\underline{s}$. The new constants used as abbreviations in what follows should be obvious.

Examples of (A) are:

(.1) (a) *The Iliad* is a story.
$Story(i)$

(.2) (a) Homer authored *The Iliad*.
Ahi

(.3) Homer authored *The Iliad* before Dostoyevsky authored *Crime and Punishment*.
$Ahi < Adc$

(.4) (a) Porphyry is an original character of *Crime and Punishment*.
$OriginalCharacterOf(p_c, c)$

(.5) (a) Porphyry is a fictional character.
$Fictional(p_c)$

(.6) (a) Porphyry is a fictional detective.
$Fictional\text{-}D(p_c)$

(.7) (a) London is a character of *A Study in Scarlet*.
$Character(l,s)$

(b) London isn't an original character of *A Study in Scarlet*.
$\neg OriginalCharacterOf(l,s)$

(c) London isn't a fictional character.
$\neg Fictional(l)$

(d) The London of *A Study in Scarlet* is an original character of *A Study in Scarlet*.
$OriginalCharacterOf(l_s,s)$

(e) The London of *A Study in Scarlet* is a fictional character.
$Fictional(l_s,s)$

(f) The London of *A Study in Scarlet* is a fictional city.
$Fictional\text{-}C(l_s)$

(g) London is not identical to the London of *A Study in Scarlet*.
$l \neq l_s$

(.8) *Crime and Punishment* is a fiction, not a false story.
$Fiction(c) \,\&\, \neg False(c)$

(.9) Sherlock Holmes is more famous than Porphyry.
Mh_sp_c

Example (.9) is interesting when compared to the superficially analogous sentence 'Sherlock Holmes is smarter than Porphyry'. For an analysis of claims like the latter, see Zalta 2000c (130–138).

(**606**) **Remark:** Analyses of the (B) and (C) Examples. We now follow a similar procedure for analyzing the (B) and (C) examples of (591).

Examples of (B) are:

(.10) (a) St. Petersburg is a city.
Cs

(b) Baker Street is in London.
Ibl

Both (a) and (b), as analyzed, are simple truths that remain true when prefaced by $c \models$ ("According to *Crime and Punishment*") and $s \models$ ("According to *A Study in Scarlet*"), respectively, as long as the names 'St. Petersburg', 'Baker Street', and 'London' are used in the storytelling to refer, in the usual way, to the well-known cities and street in Russia and England.

Examples of (C) are:

(.11) (a) In *The Iliad*, Achilles fought Hector.
$i \models Fa_ih_i$

(b) According to *Crime and Punishment*, Raskolnikov kills a pawnbroker.
$c \models \exists x(Px \,\&\, Kr_cx)$

(c) According to *A Study in Scarlet*, Holmes is a detective.
$s \models Dh_s$

(d) According to *A Study in Scarlet*, Holmes might have been a violinist.
$s \models \Diamond Vh_s$

(e) According to *Crime and Punishment*, Porphyry is a detective.
$c \models Dp_c$

(f) In *A Study in Scarlet*, London is a city.
$s \models Cl$
$s \models Cl_s$

Note that (.11.f) is ambiguous and can therefore be represented in one of two ways. In the first reading, 'London' in the data is interpreted as denoting the real city, while in the second reading, it is interpreted as denoting the fictional object: the London of *A Study in Scarlet*.

The reader may add to the foregoing any of her favorite examples that are representable in the present language of object theory. With a richer language, a wider variety of examples becomes representable.

12.6.4 Validating Judgments of Logical Consequence

(**607**) **Derivations:** Argument Representations. We now implement stage III of our methodology, as described in (591). We reproduce the arguments in (D) of (590) and, in each case, provide the formal representation of the argument as a valid derivation of object theory. It suffices, in what follows, to represent the ordinary claim 'σ is real' as $E!\sigma$:

(.12) *The Iliad* is a story.
Therefore, there are propositions that are true according to *The Iliad*.

$Story(i) \vdash \exists p(i \models p)$

(.13) *Crime and Punishment* is a story.
Therefore, something authored *Crime and Punishment*.

$Story(c) \vdash \exists x Axc$

(.14) In *Crime and Punishment*, Porphyry is a detective.
Porphyry is an original character of *Crime and Punishment*.
Therefore, Porphyry is a fictional detective.

$c \models Dp_c,\ OriginalCharacterOf(p_c, c) \vdash Fictional\text{-}D(p_c)$

(.15) London is a real city.
Therefore, London is not a fictional city.

$[\lambda x\, E!x \,\&\, Cx]l \vdash \neg Fictional(l)$

(.16) London is a real city.
London is a character of *A Study in Scarlet*.
A Study in Scarlet is a fiction.
Therefore, some characters of a fiction are real.

$[\lambda x\, E!x \,\&\, Cx]l,\ CharacterOf(l, s_1),\ Fiction(s_1) \vdash$
$\exists x \exists s(CharacterOf(x, s) \,\&\, Fiction(s) \,\&\, E!x)$

(.17) The London of *A Study in Scarlet* is an original character of *A Study in Scarlet*.
In *A Study in Scarlet*, London is a city.
Therefore, the London of *A Study in Scarlet* is a fictional city.

$OriginalCharacterOf(l_s, s),\ s \models Cl_s \vdash Fictional\text{-}C(l_s)$

(.18) London is a real city.
The London of *A Study in Scarlet* is a fictional city.
Therefore, London is not (identical to) the London of *A Study in Scarlet*.

$[\lambda x\, E!x \,\&\, Cx]l,\ Fictional\text{-}C(l_s) \vdash l \neq l_s$

(.19) Augustus Caesar worshipped Jupiter.
Jupiter is a fictional character.
Fictional characters aren't real.
Therefore, Augustus Caesar worshipped an object that isn't real.[323]

[323] For this example, we may assume that 'Jupiter' is a name derived from a text cataloguing Roman myths (though they wouldn't necessarily be thought myths at the time), e.g., Ovid's *Metamorphoses*. But, in the representation, we suppress the index to '*m*'.

$$Wcj,\ Fictional(j),\ \forall x(Fictional(x) \rightarrow \neg E!x) \vdash \exists x(Wcx \,\&\, \neg E!x)$$

(.20) Sherlock Holmes inspires some criminologists.
Sherlock Holmes is fictional.
Therefore, a fictional object inspires some criminologists.

$$\exists x(Cx \,\&\, Ih_s x),\ Fictional(h_s) \vdash \exists x \exists y(Fictional(x) \,\&\, Cy \,\&\, Ixy)$$

Notice that in (.17), the second premise is disambiguated by interpreting 'London' as denoting the London of *A Study in Scarlet* (l_s); the representation of this premise is the second one given in (606.11.f). That's because the first premise provides a context that makes it clear we are to read 'London' as denoting a fictional character.

We leave the derivations, in each case, as a simple exercise.

(**608**) **Derivations:** Theoretical Consequences of Principles. The foregoing derivations are all representations of arguments drawn from ordinary language. We now examine some derivations in which *theoretical* conclusions are drawn from the premises:

(.1) In *A Study in Scarlet*, Holmes is a detective.
Holmes is an original character of *A Study in Scarlet*.
Therefore, Holmes encodes being a detective.

$$s \models Dh_s,\ OriginalCharacterOf(h_s, s) \vdash h_s D$$

A proof is in the Appendix. Recall our assumption that the story operator is closed under relevant entailment and that this implies β-Conversion holds within its scope. Then, we may infer from $s \models Rab$ that both $s \models [\lambda x\, Rxb]a$ and $s \models [\lambda x\, Rax]b$. Then the following derivation is valid, in which the conclusion is a binary encoding claim of the form xyR:

(.2) In *The Iliad*, Achilles fought Hector.
Achilles is an original character of *The Iliad*.
Hector is an original character of *The Iliad*.
Therefore, Achilles and Hector encode the relation of fighting.

$$s \models Fh_s w_s,\ OriginalCharacterOf(h_s, s), OriginalCharacterOf(w_s, s) \vdash h_s w_s F$$

We conclude this subsection by mentioning that binary encoding claims of the form xyR will also play a role in the analysis of theoretical mathematics and will be used again in Chapter 15.

12.6.5 Final Issues Concerning Fictional Individuals

The theory of fiction raises a host of thorny philosophical problems. These problems affect every theory of fiction:

- How should we identify a fictional character when that fictional character reappears in a series of novels?

- How should we identify what seem to be distinct fictional characters that are indiscernible within a story? (See, for example, the discussion in footnote 322.)

- How can we draw comparisons among fictional characters? (E.g., While the claim 'Holmes is more famous than Poirot' was relatively easy to analyze, as in (605.9), it is not as easy to analyze 'Holmes is taller than Poirot'. Of course, one might think this last claim is false, but as long as it is deemed meaningful, one might be interested in how should it be analyzed.)

- In what sense are fictions created, if abstract objects exist necessarily and always? How can we say that the fictional object Holmes didn't exist until Conan Doyle wrote *A Study in Scarlet*?

Many philosophers have strong intuitions about these problems and especially about the facts concerning these cases; they often judge that some theory of fiction just can't be correct since it is inconsistent with what their intuitions tell them are the facts of the case. My guess is that in most of these cases, our intuitions are insufficient and just not fine-grained enough for us to put much stock in them. Rather, the way forward is to draw distinctions, formulate definitions, and test whether the theoretical conclusions we can derive from them lead us to a more refined understanding of the problem at hand.

Although I will not pursue the bulleted issues here, the following (recommendations, in some cases) should be kept in mind.

- We should always keep in mind that our primary goal is to analyze specific judgments and inferences. The data comes in the form of specific judgments and inferences made by individuals. To apply our theory, we don't need there to be some fixed body of judgments that are known and upon which everyone agrees. It suffices to analyze specific judgments and inferences by specific individuals in terms of the axioms, definitions, and theorems given above.

- Sometimes (for example, in the case where there is an ongoing series of novels with common characters and in the interval between the release of two books in the series, is a person S makes a judgment or inference about a fictional character), one has to fix the context of utterance for the judgment being analyzed, fill in the context by enquiring about additional judgments and inferences S would accept (e.g., would they accept that some proposition p is relevantly entailed by the story in question, etc.), and then use the definitions provided above to analyze the judgment or inference in question.

- As mentioned in footnote 321, authors change the expressive power of the language when they author a fiction. This has quite a number of consequences for the analysis of fiction.

- The notion of a fictional character depends on the notion of a fictional story. While an author is telling or writing down a story for the first time, they aren't referring to their characters, but only pretend to refer. Reference doesn't take place until the initial storytelling is complete. During the initial storytelling, we can appeal to the Fregean sense of the names used in the storytelling to understand the content of the author's thoughts. But we have to allow for the sense of a fictional name to change from time to time during a storytelling – this is not to say that the sense is an object that changes, but rather that at different times, different abstract objects may serve as the sense of a name.

- As implied by (591.2), being authored by something ($[\lambda x\, \exists y Ayx]$) is a contingent property of x, and the claim $Story(\sigma)$ (where σ is the name of some story), if assumed as an axiom or hypothesis, has to be regarded as modally fragile (595).

- The fact that any abstract object satisfies the definition of a story is a *contingent* fact and so the fact that Holmes satisfies the definition of a *fictional character* is a contingent fact.

- A formal system that allows one to derive suitably flagged necessary truths that are *derived from* a contingency is extremely important if to avoid the charge that contingent entities (e.g., dependent abstracta) have been misrepresented as necessary entities that have always and will always exist. (For more discussion of how our proof system for flagging necessities derived from contingencies helps Leibniz to avoid the charge Arnauld makes against the containment theory of truth, see (695) in Chapter 13.)

So of these points will be obvious from the foregoing; others have been explored in a limited way in Zalta 2000c, and Bueno & Zalta 2017. But I'm going to leave these issues and recommendations for solving them for another occasion, since further discussion would draw us too far afield. I suspect that object theory, at the very least, gives us a precise framework for drawing distinctions, formulating definitions, and testing theoretical conclusions.

(**609**) **Theorems:** Fictional Characters Aren't Possible Objects.
Theorem: A fictional object is not identical with any possible object:

$$Fictional(x) \rightarrow \neg\exists y(\Diamond E!y \; \& \; y = x)$$

Cf. Kripke 1972 [1980, 158].[324] So given that Holmes is fictional, there is a clear and provable sense in which he couldn't have been identical to any possible object.
Thm: Fictional objects are abstract.

$$Fictional(x) \rightarrow A!x$$

Cf. Kripke 1973 [2013, 73].[325]

(**610**) **Remark:** What It Means To Say that A Character Might Exist. Suppose fictional character x originates in some story, say $\underline{s}$. Then x is abstract and encodes the properties attributed to x in $\underline{s}$. However, we commonly suppose that x, considered as a fictional character, might have existed (or might have been real). For example, there is a sense in which the character, Sherlock Holmes, might have existed. As might be expected, we can't regard the English "might have existed" as saying that the object Holmes might have been a concrete object. Holmes can't be identified as a possibly concrete object since there are too many of those that are consistent with the stories. The proper name 'Holmes' should be given a theoretically-described denotation, but since there are innumerable properties that aren't attributed to Holmes in the novels, we have no theoretical means of picking out a distinctive possibly concrete object. At first glance, however, it seems natural to interpret the claim "Holmes might have existed" as "Possibly, something exemplifies all of the properties Holmes exemplifies in the stories". This would imply that, at some possible world w,

[324] In Kripke 1972 [1980, 158], we find:

> Similarly, I hold the metaphysical view that, granted that there is no Sherlock Holmes, one cannot say of any possible person that he would have been Sherlock Holmes, had he existed. Several distinct possible people, and even actual ones such as Darwin or Jack the Ripper, might have performed the exploits of Holmes, but there is none of whom we can say that he would have been Holmes had he performed these exploits. For if so, which one? (See my 'Semantical Considerations on Modal Logic', ...) The quoted assertion gives the erroneous impression that a fictional name such as 'Holmes' names a particular possible-but-not-actual individual.

[325] In Kripke 1973 [2013, 73], we find:

> A fictional character, then, is in some sense an abstract entity. It exists in virtue of more concrete activities of telling stories, writing plays, writing novels and so on, under criteria which I won't try to state precisely, but which should have their own obvious intuitive character. It is an abstract entity which exists in virtue of more concrete activities the same way that a nation is an abstract entity which exists in virtue of concrete relations between people.

Note that you can't instantiate the theorems and definitions with examples until there is some human activity: a storytelling! This accounts for the sense (recall Kripke) in which they exist in virtue of human activity. Conan Doyle enriched the expressive power of language by using 'Holmes', 'Watson', 'Moriarty' in a storytelling. See Zalta 2006a.

there is a concrete object y that exemplifies at w the properties attributed to x in $\underline{s}$.

But an issue arises for this interpretation. Consider the fact that in *The Iliad*, Achilles fought Hector. So to suppose that Achilles might have existed, under this interpretation, would be to suppose that, at some possible world w, an object y exemplifies at w the properties attributed to Achilles in *The Iliad*. One of those properties is *fighting Hector*. But we've also identified Hector as an abstract object (given that he, too, originated in *The Iliad*). So if Achille exemplifies *fighting Hector* at w, it would follow that there is an abstract object z that is identical to Hector and such that at w, y exemplifies *fighting* z.

This result seems rather puzzling, since *fighting*, unlike *loves*, *inspires*, *worships*, etc., is not an intensional relation. Necessarily, if x fights y, then both $E!x$ and $E!y$. What this shows is we can't rest with this simple analysis of what it is to say that a character might have existed. Rather, to say the fictional character x of story $\underline{s}$ might have existed is to say that all the fictional characters of $\underline{s}$ might have jointly existed, in the sense that, for any fictional characters $x_1,\ldots,x_n$ of story $\underline{s}$, there is a possible world w and objects $y_1,\ldots,y_n$ such that each of the $y_1,\ldots,y_n$ exemplify at w, respectively, all of the properties attributed to $x_1,\ldots,x_n$.

While this seems to be a step closer to what we want, it still leaves the previous problem. If to say that Achilles might have existed is to say that all of the characters of the Iliad might have existed in the sense just indicated, then *a fortiori* there has to be a possible world w, and objects y_1 and y_2 such that y_1 exemplifies all of the properties attributed to Achille in the *Iliad* and y_2 exemplifies all of the properties attributed to Hector in the *Iliad*. But, without loss of generality, note that the property of *fighting Hector* is attributed to Achille in the *Iliad*. Yet we want y_1 to exemplify at w the property of *fighting* y_2, not the property of *fighting Hector*.

I think the best way of understand the issue here is to suppose that our intuitions about Achille's possible existence should be analyzed by *going through the language of the fiction*. That is, in order to formalize the intuition that a fictional story $\underline{s}$ is true at some world w, we have to suppose that for each individual term κ that designates an original character of $\underline{s}$, there is an ordinary object y such that for any formula φ, if φ is true in $\underline{s}$, then φ' is true at w, where φ' is the result of substituting y for κ in φ. When this metalinguistic fact holds for $\underline{s}$ and w, we say that $\underline{s}$ is true at w. Then we may say that a fictional story $\underline{s}$ is *possibly true* just in case there is some possible world w where $\underline{s}$ is true at w. Finally, we may say that an original character x of story $\underline{s}$ *might have existed* if and only if s is possibly true. Thus, facts about language are needed to analyze the sense in which a fictional character might have existed.

Formally, we can express the foregoing analysis using the following *metathe-*

oretical definitions, for any finite number n of individual terms κ occurring in the story ($n \geq 1$):

- A fictional story $\underline{s}$ *is true at* w if and only if $\forall\kappa_1 \ldots \forall\kappa_n$
  $\big((OriginalCharacterOf(\kappa_1, \underline{s})\ \&\ \ldots\ \&\ OriginalCharacterOf(\kappa_n, \underline{s})) \rightarrow$
  $\exists y_1 \ldots \exists y_n \forall\varphi(\underline{s} \models \varphi \rightarrow w \models \varphi')\big)$,
  where φ' is the result of substituting y_i for κ_i in φ, provided y_i is substitutable for κ_i ($1 \leq i \leq n$).

- A fictional story $\underline{s}$ is *possibly true* if and only if $\exists w(\underline{s}$ is true at $w)$.

- x *might have existed* if and only if $\exists\underline{s}(OriginalCharacter(x, s)\ \&\ \underline{s}$ is possibly true)

Of course, these definitions will have to be adjusted to account for the truth of fictional stories that involve fictional properties as well as fictional individuals, so that we can define what it means to say that a fictional property might have existed. But we can leave this consideration for the section of Chapter 15 where we analyze fictional properties and fictional relations in typed object theory.

Chapter 13

Concepts

It is not always clear what is meant when philosophers talk about *concepts*. In this chapter, we define a notion of *concept* governed by a variety of interesting theorems, many of which represent principles that intuitively characterize this notion. Furthermore, some of the theorems that we derive look very similar to principles that Leibniz adopted in his work. Though many philosophers have supposed that Leibnizian concepts, like Plato's Forms, are to be analyzed as properties, the theorems below establish that Leibnizian principles about concepts fall out very naturally when the latter are analyzed as abstract individuals. Indeed, in (678) and (735) below, we explain difficulties that would arise if one were to analyze Leibnizian concepts as properties.

Thus, in what follows Leibnizian concepts are to be distinguished from Fregean concepts, which *are* taken to be properties in Chapter 14 (see the opening lines of Section 14.2, where this is justified). But whether or not one agrees that the following theorems constitute a good interpretation of Leibniz, we take it that they constitute an interesting and compelling philosophical theory of concepts in their own right.

(**611**) **Remark:** Leibnizian Concepts. Over the course of his life, Leibniz developed three different strands of his theory of concepts:

- a non-modal '*calculus*' of concepts,
- a *concept containment* theory of truth, and
- a *modal metaphysics* of complete, individual concepts.

We discuss these in turn.

To modern eyes, Leibniz's so-called 'calculus' of concepts is really an algebra, though Leibniz probably used 'calculus' in the sense of *calculus ratiocinator*, a general framework for reasoning. He produced fragments of this algebra throughout his life, but only in his late works (1690a, 1690b) did he explicitly introduce a primitive operation symbol (which we shall write as $\oplus$) so that he could write $A \oplus B$ to denote the *sum* of concepts A and B.[326] He also introduced the notions of concept *containment* and concept *inclusion*, so that he could say that the concept $A \oplus B$ contains both the concepts A and B, and that both A and B are included in $A \oplus B$. Some of the key principles that emerge from this fragment of Leibniz's work are: that concept addition is idempotent, commutative, and associative; that concept containment and inclusion are reflexive, anti-symmetric, and transitive; that if concept A is included in concept B, then there is a concept C such that $A \oplus C = B$; and that if A is included in B, then $A \oplus B = B$.

Leibniz advocated a concept containment theory of truth throughout his life, though he often expressed it in terms of concept inclusion. Here is a classic statement, from his correspondence with Arnauld (June 1686, LA 63, G.ii 56):

> ... in every true affirmative proposition, necessary or contingent, universal or particular, the concept of the predicate is in a sense included in that of the subject; the predicate is present in the subject.

Stated in terms of containment, this becomes the claim that in a true subject-predicate statement, the concept of the subject contains the concept of the predicate. In what follows, we shall analyze the concept of a predicate in the material mode, as the concept of a *property*. Moreover, we shall define the concepts of such subjects as 'Alexander' and 'every person' as they occur in the sentences 'Alexander is rational' and 'Every person is rational'. We show that from the modern analysis of the claim that Alexander is rational, one can derive our analysis of the Leibnizian claim that the concept of Alexander contains the concept of *being rational*. Moreover, we then show that from the modern analysis of the claim that every person is rational, one can derive our analysis of the Leibnizian claim that the concept *every person* contains the concept of *being rational*.

In middle and late period works (*Discourse on Metaphysics* (1686), *Theodicy* (1709), and *The Monadology* (1714)), Leibniz developed a modal metaphysics of individual concepts that he used to analyze modal facts about individuals in terms of facts about various *individual concepts* that *appear at* other possible worlds. Some fundamental principles underlying Leibniz's view are:

- If an ordinary individual u is F but might not have been F, then (i) the individual concept of u contains the concept of F, and (ii) there is an individual concept that: (a) is a counterpart of the concept of u, (b) fails to contain the concept of F, and (c) appears at some other possible world.

[326] Prior to those works, Leibniz indicated concept addition by concatenating the symbols for two concepts.

- If an ordinary individual u isn't F but might have been F, then (i) the individual concept of u fails to contain the concept of F, and (ii) there is an individual concept that: (a) is a counterpart of the concept of u, (b) contains the concept of F, and (c) appears at some other possible world.

Although Leibniz never actually states these principles explicitly, they are implicit in the views he expressed. As part of our goal of analyzing and unifying the various components of Leibniz's theory of concepts, we shall formalize and prove the above fundamental principles of Leibniz's modal metaphysics of concepts. The formalizations and derivations that we construct in analyzing Leibniz's theory realize, at least in part, his idea of *calculus ratiocinator*, though we won't argue for this here.

The developments that follow revise and enhance the work in Zalta 2000a in numerous ways. The discussion and procession of theorems has been revised, and the statement of the theorems and their proofs have been improved. Though a comparison of our work below with other work in the secondary literature on Leibniz would be useful, it is not attempted here. Note, however, that some commentators treat only Leibniz's nonmodal calculus of concepts and not the modal metaphysics of individual concepts,[327] while others treat only the modal metaphysics of individual concepts and not the nonmodal calculus.[328] Although Lenzen (1990) treats both, his work provides us with a *model* of Leibnizian concepts within set theory; it does not provide a *theory* of concepts. By contrast, the following reconstruction makes no set-theoretic assumptions, unlike most of the works in the secondary literature just cited. As mentioned previously, the work has significant philosophical interest in its own right, whether or not it constitutes a contribution to the secondary literature on Leibniz. That's why 'Leibnizian' is sometimes used within parentheses.

(**612**) **Definition:** (Leibnizian) Concepts. The key idea underlying our analysis of (Leibnizian) concepts is that they are abstract individuals. Consequently, we identify the property *being a concept* with the property *being abstract*:

$$C! =_{df} A!$$

Thus, all of the previous theorems about abstract individuals become theorems about (Leibnizian) concepts. Note also that in light of theorem (180.2), $C!x \rightarrow \Box C!x$. Hence, by GEN, we know $\vdash_{\Box} \forall x(C!x \rightarrow \Box C!x)$. So $C!x$ is a rigid condition, in the sense of (260.1). We therefore leave it as an exercise to show that $C!x$ is a rigid restriction condition in the sense of (340).

[327] See, for example, Rescher 1954; Kauppi 1960, 1967; Castañeda 1976, 1990; and Swoyer 1994, 1995.

[328] See Mates 1968, Mondadori 1973, and Fitch 1979.

(**613**) **Theorems:** Immediate Equivalences and Identities. It is now a simple consequence of the previous definition that x is a concept that encodes exactly the properties such that φ if and only if x is an abstract object that encodes exactly the properties such that φ:

$$(C!x \,\&\, \forall F(xF \equiv \varphi)) \equiv (A!x \,\&\, \forall F(xF \equiv \varphi))$$

(**614**) **Theorems:** Concept Comprehension and Canonical Concept Descriptions. From the preceding theorems, we can easily prove that (.1) there is a concept that encodes exactly the properties F such that φ; (.2) there is a unique concept that encodes exactly the properties F such that φ; (.3) the individual that is a concept encoding exactly the properties such that φ exists; and (.4) the concept that encodes exactly the properties such that φ is identical to the abstract object that encodes exactly the properties such that φ:

(.1) $\exists x(C!x \,\&\, \forall F(xF \equiv \varphi))$, provided x doesn't occur free in φ

(.2) $\exists!x(C!x \,\&\, \forall F(xF \equiv \varphi))$, provided x doesn't occur free in φ

(.3) $\imath x(C!x \,\&\, \forall F(xF \equiv \varphi))\downarrow$, provided x doesn't occur free in φ

(.4) $\imath x(C!x \,\&\, \forall F(xF \equiv \varphi)) = \imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$,
provided x doesn't occur free in φ

By (.3), any appropriate φ can be used to formulate a *canonical concept description* of the form $\imath x(C!x \,\&\, \forall F(xF \equiv \varphi))$ and, by (.4), the concept so described is identical to a canonical abstract object. Henceforth all of the machinery for, and discussions of, canonical and strictly canonical descriptions, in (253) – (262), can be repurposed for concepts. These facts come into play on numerous occasions below.

(**615**) **Remark:** Restricted Variables and Canonical Concepts. Since we saw, at the end of (612), that $C!x$ is a rigid restriction condition, we may introduce the variables $c, d, e, f, \ldots$ as rigid restricted variables ranging over concepts (thus, we are repurposing c from its use in Chapter 10 as a non-rigid restricted variable ranging over classes). So we assert theorems, and reason, with $c, d, \ldots$ as free restricted variables, and employ our extended Rule RN (341).

We shall also use $c_1, c_2, \ldots$, $d_1, d_2, \ldots$, etc., as primitive constants introduced to pick out an arbitrary concept. Since it is clear that concepts exist, it follows that $\forall c\varphi \rightarrow \exists c\varphi$ and we hereafter assume it; cf. Remark (342). Moreover, when κ is any restricted variable or primitive constant for an arbitrary concept, we may assume $\kappa\downarrow$ is axiomatic (39.2), that κ can be instantiated into any universal claim, and that $\kappa = \kappa$ can be introduced by Rule =I.

With our rigid restricted variables for concepts, we may express (614.1) – (614.4) as follows, provided c doesn't occur free in φ:

$$\exists c\forall F(cF \equiv \varphi) \tag{614.1}$$

$$\exists! c\forall F(cF \equiv \varphi) \tag{614.2}$$

$$\imath c\forall F(cF \equiv \varphi)\downarrow \tag{614.3}$$

$$\imath c\forall F(cF \equiv \varphi) = \imath x(A!x \,\&\, \forall F(xF \equiv \varphi)) \tag{614.4}$$

By a modest abuse of language, we may say that the description $\imath c\forall F(cF \equiv \varphi)$ describes a *canonical concept*. Since canonical concepts are canonical abstract objects, the theorems governing the latter apply to the former. For example, it is straightforward to show that a version of the Abstraction Principle (256.2)★ applies to concepts, namely, the concept encoding all and only the properties such that φ encodes F if and only if φ:

$\imath c\forall F(cF \equiv \varphi))F \equiv \varphi$, provided c doesn't occur free in φ

For certain formulas φ, we'll be able to establish that the canonical concept is strictly canonical and so subject to a modally strict version of the above theorem schema.

(**616**) **Theorems:** Identity. In LLP 131–132 (G.vii 236), Propositions 1 and 3, Leibniz uses the substitution of identicals to derive that the notion of identity, as it applies to concepts, is symmetrical and transitive. However, '=' is defined in the present theory. Given theorems (117.1) – (117.3), it immediately follows that identity is reflexive, symmetrical, and transitive on the concepts. Hence, using our restricted variables, we have:

(.1) $c = c$

(.2) $c = d \rightarrow d = c$

(.3) $c = d \,\&\, d = e \rightarrow c = e$

Note that in Definition 1 of G.vii 236, Leibniz uses '∞' as the *identity* symbol. In LLP 131, the translation of of G.vii 236 includes the claims "'A = B' means that A and B are the same" and that "those terms are 'the same' ... of which either can be substituted for the other wherever we please without loss of truth".

13.1 The Calculus of Concepts

13.1.1 Concept Addition

(**617**) **Definition:** Concept Addition (Summation). To define concept addition or summation, let us say that concept c is a *sum of* concepts *d and e* if and only if c encodes all and only the properties encoded by either d or e:

$$SumOf(c,d,e) \equiv_{df} \forall F(cF \equiv dF \vee eF)$$

To produce an example of our definition using modally strict reasoning, note that the following are instances of concept comprehension (614.1), where P, Q, and R are any arbitrarily chosen properties:

- $\exists c \forall F(cF \equiv F = P)$
- $\exists c \forall F(cF \equiv F = Q \vee F = R)$
- $\exists c \forall F(cF \equiv F = P \vee F = Q \vee F = R)$

Let c_1, c_2, and c_3, respectively, be arbitrary such objects. Since c_1 encodes just the property P, c_2 encodes just the properties Q and R, and c_3 encodes just the properties P, Q, and R, it is an easy exercise to show $SumOf(c_3,c_1,c_2)$.

(**618**) **Theorems:** The Sum of Concepts d and e Exists. In the usual manner, it follows that (.1) concepts d and e have a sum; (.2) concepts d and e have a unique sum; and (.3) the sum of concepts d and e exists:

(.1) $\exists c SumOf(c,d,e)$

(.2) $\exists ! c SumOf(c,d,e)$

(.3) $\imath c SumOf(c,d,e)\downarrow$

It is worth mentioning that since d and e are restricted variables, these theorems are really shorthand for *conditional* existence claims. For example, (.1) is shorthand for $(C!x \,\&\, C!y) \rightarrow \exists z(C!z \,\&\, SumOf(z,x,y))$. But since we know that concepts exist, we can derive unconditional existence claims from them.

(**619**) **Definition:** The Sum of Concepts d and e. By our last theorem, we are entitled to introduce the following notation for *the* concept that is the sum of concepts d and e:

$$d \oplus e \;=_{df}\; \imath c SumOf(c,d,e)$$

Since there are free restricted variables occurring in the definition, it is important to remember the discussion in (339), in which we discuss the inferential role of definitions that employ free restricted variables. Given that discussion and the modally strict fact that $C!x \rightarrow \Box C!x$ (mentioned previously in (612)), it should be clear that an expression of the form $\kappa \oplus \kappa'$ is a binary functional term that is significant only when it is known, either by proof or by hypothesis, that $C!\kappa$ and $C!\kappa'$. Moreover, since concepts are just abstract individuals, it should also be clear that $d \oplus e$ exists whenever d and e are any abstract objects.

(**620**) **Lemmas:** Strict Canonicity of Sums. In the usual way, it is straightforward to show that $d \oplus e$ is (identical to) a canonical concept:

(.1) $d \oplus e = \imath c \forall F(cF \equiv dF \vee eF)$

If we let φ be the formula $dF \vee eF$, then we may use definition (260.1) to establish, by modally strict means, the following universal claim, which therefore tells us (.2) φ is a rigid condition on properties:

(.2) $\forall F((dF \vee eF) \rightarrow \Box(dF \vee eF))$

Hence, $d{\oplus}e$ is a *strictly* canonical concept, by (260.2). By theorem (261.2) and the definition of $C!$, it follows that (.3) the sum of d and e is a concept that encodes all and only the properties F such that either d encodes F or e encodes F:

(.3) $C!d{\oplus}e \ \&\ \forall F(d{\oplus}eF \equiv dF \vee eF)$

Note that the first conjunct of (.3) is an exemplification formula consisting of a unary relation term $C!$ and a complex individual term $d{\oplus}e$. Moreover, it clearly follows from (.3) by definition (617) that (.4) $d{\oplus}e$ is a sum of d and e:

(.4) $SumOf(d{\oplus}e, d, e)$

(.1) – (.4) are modally strict theorems used frequently in the proofs of subsequent theorems.

(**621**) **Theorems:** Concept Addition Forms a Semi-Lattice. It follows straightforwardly from the previous lemma that $\oplus$ is idempotent, commutative, and associative:

(.1) $c{\oplus}c = c$

(.2) $c{\oplus}d = d{\oplus}c$

(.3) $(c{\oplus}d){\oplus}e = c{\oplus}(d{\oplus}e)$

In virtue of the last fact, we may leave off the parentheses in the expressions $(c \oplus d) \oplus e$ and $c \oplus (d \oplus e)$.

Thus, concept addition behaves in the manner that Leibniz prescribed. He took the first two of these theorems as *axioms* of his calculus, whereas we derive them as theorems.[329] Unfortunately, he omitted associativity from his list of axioms for $\oplus$; as Swoyer (1995, 1994) points out, it must be included for the proofs of certain theorems to go through.

(**622**) **Remark:** Concept Addition and Properties. We've analyzed concepts as abstract individuals and concept addition as a functional condition on concepts. Of course, Leibniz's texts use variables for concepts that suggest that he

[329] See LLP 132 (G.vii 237), Axioms 2 and 1, respectively. Other idempotency assertions appear in LLP 40 (G.vii 222), LLP 56 (C 366), LLP 85 (C 396), LLP 90 (C 235), LLP 93 (C 421), and LLP 124 (G.vii 230). Swoyer (1995, footnote 5) also cites C 260 and C 262. Lenzen (1990) cites GI 171 for idempotency. Other commutativity assertions appear in LLP 40 (G.vii 222), LLP 90 (C 235), and LLP 93 (C 421).

conceived of concepts as properties. For example, in 1690a and 1690b, he uses variables $A, B, \ldots, L, M, N$, etc., and he instantiates these variables with predicative expressions such as 'triangle', 'trilateral', 'rational', 'animal', etc. One may certainly try to reconstruct Leibnizian concepts as properties, but we now discuss an issue that arises for that analysis. On the property-theoretic analysis, concept addition is traditionally regarded as property conjunction, so that one would define:

$$F + G =_{df} [\lambda x\, Fx \,\&\, Gx]$$

So to derive (621.1) – (621.3) as theorems, one would have to derive:

$$[\lambda x\, Fx \,\&\, Fx] = F$$

$$[\lambda x\, Fx \,\&\, Gx] = [\lambda x\, Gx \,\&\, Fx]$$

$$[\lambda x\, [\lambda y\, Fy \,\&\, Gy]x \,\&\, Hx] = [\lambda x\, Fx \,\&\, [\lambda y\, Gy \,\&\, Hy]x]$$

If one were to take properties to be identical when materially (i.e., extensionally) equivalent, the above would be easy consequences. But such a definition of property identity is incorrect. Similarly, if one were to take properties to be identical when necessarily equivalent (i.e., by defining $F = G$ as $\Box\forall x(Fx \equiv Gx)$), one would also obtain the above consequences. But, again, this definition is incorrect and the present theory doesn't endorse it. We have formulated our system so that one can consistently assert that there are properties F and G such that both $\Box\forall x(Fx \equiv Gx)$ and $F \neq G$.

Thus, if one allows for distinct properties that are necessarily equivalent with respect to exemplification, then it is not clear how to derive the above property identities in absence of further axioms governing property identity. Our policy has been to eschew such axioms. The present theory offers precise conditions under which properties exist and precise conditions under which they are identical. The latter tell us exactly what we are asserting or need to prove when we assert or prove that properties F and G are either identical or distinct. But, we leave it as an open question, to be decided as the case may demand, whether the property identities displayed above are to be endorsed or not.

By understanding Leibnizian concepts as abstract individuals, one can derive (621.1) – (621.3) as theorems, but if one prefers the analysis of Leibnizian concepts as properties and takes properties seriously as intensional entities, then more work has to be done to derive the identities needed to show that property addition, as defined in this Remark, is idempotent, commutative, and associative. Moreover, it should also be noted that once *concepts of properties* are introduced in Section 13.2, where we define *the concept of F* ($\boldsymbol{c}_F$), then the following instances of (621.1) – (621.3) bring us even closer to Leibniz's texts:

$$\boldsymbol{c}_F \oplus \boldsymbol{c}_F = \boldsymbol{c}_F$$

$$\boldsymbol{c}_F \oplus \boldsymbol{c}_G = \boldsymbol{c}_G \oplus \boldsymbol{c}_F$$

$$(\boldsymbol{c}_F \oplus \boldsymbol{c}_G) \oplus \boldsymbol{c}_H = \boldsymbol{c}_F \oplus (\boldsymbol{c}_G \oplus \boldsymbol{c}_H)$$

These results at least provide an interpretation of the Leibnizian texts in which he asserts concept addition is idempotent and commutative (and assumes associativity).

(**623**) **Theorems:** Concept Addition and Identity. Leibniz proves two other theorems pertaining solely to concept addition and identity in LLP 133–4 (G.vii 238), Propositions 9 and 10:

(.1) $c = d \rightarrow c \oplus e = d \oplus e$

(.2) $c = d \;\&\; e = f \rightarrow c \oplus e = d \oplus f$

In the notes following Propositions 9 and 10 in LLP 133-134 (G.vii 238), Leibniz observes that counterexamples to the converses of these theorems can be produced. To produce a counterexample to the converse of (.1), first let P, Q, and R be any three distinct properties. We know there are such by (225.7). Then:

- let c_1 be $\imath c \forall F(cF \equiv F = P \vee F = Q \vee F = R)$
- let c_2 be $\imath c \forall F(cF \equiv F = P \vee F = Q)$
- let c_3 be $\imath c \forall F(cF \equiv F = R)$

Then it is straightforward to show that $c_1 \oplus c_3 = c_2 \oplus c_3$ and $c_1 \neq c_2$, contrary to the converse of (.1).

Similarly, to produce a counterexample to the converse of (.2), first let P, Q, R, and S be any four distinct properties. We know there are such by (225.7). Then:

- let c_1 be $\imath c \forall F(cF \equiv F = P \vee F = Q)$
- let c_2 be $\imath c \forall F(cF \equiv F = P)$
- let c_3 be $\imath c \forall F(cF \equiv F = R \vee F = S)$
- let c_4 be $\imath c \forall F(cF \equiv F = Q \vee F = R \vee F = S)$

It is then easy to show that $c_1 \oplus c_3 = c_2 \oplus c_4$ and $c_1 \neq c_2$ (and, indeed, $c_3 \neq c_4$), contrary to the converse of (.2).

13.1.2 Concept Inclusion and Containment

(**624**) **Definitions:** Inclusion and Containment. It is an algebraic fact that an idempotent, commutative, and associative operation on a domain induces a partial ordering on that domain. In the present case, concept addition induces the partial ordering of *concept inclusion* ($c \preceq d$) and a converse ordering of *concept containment* ($d \succeq c$). We'll prove these facts below. Leibniz was aware of this connection and derives these facts for the case of concept inclusion. In Definition 3 of LLP 132 (G.vii 237), Leibniz defined both $c \preceq d$ and $d \succeq c$ as $\exists e(c \oplus e = d)$.[330] Moreover, in Propositions 13 and 14 of LLP 135 (G.vii 239), Leibniz derives both directions of the equivalence $c \preceq d \equiv c{\oplus}d = d$ as theorems.

In object theory, however, a deeper level of analysis of concept inclusion and concept containment is available. Once that analysis is formulated, it follows that inclusion and containment partially order the concepts; this is item (625) below. Moreover, Leibniz's definition of inclusion in Definition 3 and the equivalence that he obtains via his Propositions 13 and 14 are both derivable from that analysis. These are theorems (628) and (629) below.

We begin by defining: *c is included in d* just in case *d* encodes every property *c* encodes. Formally:

(.1) $c \preceq d \ \equiv_{df} \ \forall F(cF \rightarrow dF)$

We shall see, in what follows, that this notion of concept inclusion is a generalization of the notion of *part-of*, which was defined on situations in (475).

Leibniz's notion of concept containment is now just the converse of inclusion. Let us say that *d contains c* just in case *c* is included in *d*:

(.2) $d \succeq c \ \equiv_{df} \ c \preceq d$

Consequently, the theorems below are developed in pairs: one member of the pair governs concept inclusion and the other concept containment. However, in the Appendix, we prove the theorem only as it pertains to concept inclusion.

(**625**) **Theorems:** Concept Inclusion and Containment Are Partial Orders. It now follows that concept inclusion and containment are reflexive, anti-symmetric, and transitive. To show anti-symmetry, we use $c \npreceq d$ to abbreviate $\neg(c \preceq d)$, and $c \nsucceq d$ to abbreviate $\neg(c \succeq d)$:

[330] Strictly speaking, Leibniz didn't use the existential quantifier in his definition of concept inclusion and containment. His definition reads as follows:

> Definition 3. That *A* 'is in' *L*, or, that *L* 'contains' *A*, is the same as that *L* is assumed to be coincident with several terms taken together, among which is *A*.

This is the translation in LLP 132 of the passage in G.vii 237. It is not entirely clear to me what 'taken together' means here, but our representation of this definition may capture at least one correct way of interpreting this passage.

(.1) $c \preceq c$
$c \succeq c$

(.2) $c \preceq d \rightarrow (c \neq d \rightarrow d \npreceq c)$
$c \succeq d \rightarrow (c \neq d \rightarrow d \nsucceq c)$

(.3) $c \preceq d \;\&\; d \preceq e \rightarrow c \preceq e$
$c \succeq d \;\&\; d \succeq e \rightarrow c \succeq e$

See LLP 133 (= G.vii 238), Proposition 7, for Leibniz's proof of the reflexivity of inclusion. See LLP 135 (= G.vii 240), Proposition 15, for Leibniz's proof of the transitivity of inclusion. See also LLP 33 (= G.vii 218) for the reflexivity of containment.

(**626**) **Theorems:** Inclusion, Containment, and Identity. Leibniz proves in LLP 136 (G.vii 240), Proposition 17, that when concepts c and d are included, or contained, in each other, they are identical. Hence we have the more general theorem:

(.1) $c = d \equiv c \preceq d \;\&\; d \preceq c$
$c = d \equiv c \succeq d \;\&\; d \succeq c$

Two interesting further consequences of concept inclusion and identity are that (.2) concepts c and d are identical just in case the same concepts are included in c and d, and (.3) concepts c and d are identical just in case c and d are included in the same concepts:

(.2) $c = d \equiv \forall e(e \preceq c \equiv e \preceq d)$
$c = d \equiv \forall e(c \succeq e \equiv d \succeq e)$

(.3) $c = d \equiv \forall e(c \preceq e \equiv d \preceq e)$
$c = d \equiv \forall e(e \succeq c \equiv e \succeq d)$

(**627**) **Theorems:** Inclusion and Addition. In LLP 33 (G.vii 218), Leibniz asserts '*ab* is *a*' and '*ab* is *b*'. Here it looks as if *ab* is to be interpreted as $a \oplus b$ and 'is' as containment. (This is an application of Leibniz's containment theory of truth, which will be discussed below.) Thus, in our system, these claims become: the sum of c and d contains c, and the sum of c and d contains d. In each case, we state the inclusion version first, namely, (.1) c is included in the sum of c and d, and (.2) d is included in the sum of c and d:

(.1) $c \preceq c \oplus d$
$c \oplus d \succeq c$

(.2) $d \preceq c \oplus d$
$c \oplus d \succeq d$

Leibniz also notes that if c is included in d, then the sum of e and c is included in the sum of e and d:

(.3) $c \preceq d \;\rightarrow\; e \oplus c \preceq e \oplus d$
$c \succeq d \;\rightarrow\; e \oplus c \succeq e \oplus d$

See LLP 134 (G.vii 239), Proposition 12. See also LLP 41 (G.vii 223), for the version governing containment. Note that there is a counterexample to the converse of (.3). If we let e_1 and c_1 both be $\imath c \forall F(cF \equiv F = P)$, and let d_1 be $\imath c \forall F(cF \equiv F = Q \vee F = R)$, where P, Q, R are all pairwise distinct, then it is easy to show that $e_1 \oplus c_1 \preceq e_1 \oplus d_1$ and $c_1 \npreceq d_1$.

It also follows that $c \oplus d$ is included in e if and only if both c and d are included in e:

(.4) $c \oplus d \preceq e \;\equiv\; c \preceq e \,\&\, d \preceq e$
$e \succeq c \oplus d \;\equiv\; e \succeq c \,\&\, e \succeq d$

Leibniz notes a more economical form of the left-to-right direction of (.4). In LLP 136 (G.vii 240), Corollary to Proposition 15, he argues that if $A \oplus N$ is in B, then N is in B. However, he proves the right-to-left direction of (.4) at LLP 137 (G.vii 241), Proposition 18.

Finally, we may prove that if c is included in d and e is included in f, then $c \oplus e$ is included in $d \oplus f$:

(.5) $c \preceq d \,\&\, e \preceq f \;\rightarrow\; c \oplus e \preceq d \oplus f$
$c \succeq d \,\&\, e \succeq f \;\rightarrow\; c \oplus e \succeq d \oplus f$

See LLP 137 (= G.vii 241), Proposition 20.

Exercise. Find a counterexample to the converse direction of (.5).

13.1.3 Concept Inclusion, Addition, and Identity

The following theorems establish that our definitions of concept inclusion (containment), addition, and identity are related in the way required by Leibniz's theory of concepts.

(**628**) **Theorem:** Leibnizian Definition of Inclusion. Our efforts thus far imply that c is included in d if and only if, for some concept e, the sum of c and e identical to d:

$$c \preceq d \equiv \exists e(c \oplus e = d)$$
$$c \succeq d \equiv \exists e(c = d \oplus e)$$

If we have correctly understood Leibniz's Definition 3, in LLP 132 = G.vii 237 (see footnote 330), then the proof of the inclusion version constitutes a derivation of his definition as a theorem.

(**629**) **Theorem:** Leibniz's Equivalence. Our work now allows us derive the principal theorem of Leibniz's non-modal calculus of concepts; it governs concept identity, concept inclusion, and concept addition and asserts that c is included in d if and only if the sum of c and d is identical to d:

$$c \preceq d \equiv c \oplus d = d$$
$$c \succeq d \equiv c = c \oplus d$$

See LLP 135 (G.vii 239), Propositions 13 and 14. The notions $=$, $\oplus$, and $\preceq$ are all defined and, other than the substitution of identicals (41), the axioms and theorems Leibniz used to prove this claim have all been derived as theorems.

(**630**) **Theorem:** Leibniz's Proposition 23. In LLP 140, Proposition 23 is stated as "Given two disparate terms, A and B, to find a third term, C, different from them and such that A ⊕ B = A ⊕ C" (cf. G.vii 243). We can capture this as follows: (.1) If c is not included in d and d is not included in c, then there is a concept e such that (a) e is distinct from both c and d and (b) the sum of c and e is identical to the sum of c and d. Formally:

(.1) $(c \npreceq d \;\&\; d \npreceq c) \rightarrow \exists e(e \neq c \;\&\; e \neq d \;\&\; c \oplus e = c \oplus d)$

As the first steps of the proof, assume the antecedent and then show that $c \oplus d$ is a witness to the existential claim; since c and d each encodes a property the other doesn't, $c \oplus d$ will be distinct from both c and d.

Note also the following consequence of our definitions, namely: (.2) c is included in d and d is not included in c if and only if some concept e not included in c is such that the sum of c and e is identical to d:

(.2) $(c \preceq d \;\&\; d \npreceq c) \equiv \exists e(e \npreceq c \;\&\; c \oplus e = d)$

This allows for the degenerate case in which c encodes no properties and d encodes one or more properties.

13.1.4 The Algebra of Concepts

In this subsection we first investigate the algebraic principles derivable from our theory of concepts and then consider the extent to which this algebra constitutes a mereology. To anticipate a bit, we prove that concepts are structured not only as a bounded lattice, but also as a complete Boolean algebra bounded above (by a universal concept) and below (by a null concept). Then, in the next subsection, i.e., Section 13.1.5, we examine the extent to which our algebra of concepts obeys the principles of mereology when our notion of concept inclusion, $x \preceq y$, is interpreted as: *x is a part of y*.

(**631**) **Definition:** Concept Multiplication (i.e., Concept Products). Recall that theorem (621) tells us that addition is idempotent, commutative, and associative on concepts and so forms a semi-lattice (following standard mathematical practice). Over the course of the next several items, we show that concepts form not just a semi-lattice but a *lattice*. We define concept multiplication by saying that c is a *product of* d and e just in case c encodes all and only the properties that d and e encode in common:

$$ProductOf(c,d,e) \equiv_{df} \forall F(cF \equiv dF \,\&\, eF)$$

For example, where P, Q, R are three, pairwise-distinct properties:

let c_1 be $\imath c\forall F(cF \equiv F = P \vee F = Q)$

let c_2 be $\imath c\forall F(cF \equiv F = Q \vee F = R)$

let c_3 be $\imath c\forall F(cF \equiv F = Q)$

Since c_1 encodes just the properties P and Q, c_2 encodes just the properties Q and R, and c_3 encodes just the property Q, it is an easy exercise to show $ProductOf(c_3, c_1, c_2)$.

(**632**) **Theorems:** Existence of Products. In the usual way, we prove (.1) there exists a product of concepts d and e; (.2) there exists a unique product of concepts d and e; and (.3) the product of concepts d and e exists:

(.1) $\exists c ProductOf(c,d,e)$

(.2) $\exists! c ProductOf(c,d,e)$

(.3) $\imath c ProductOf(c,d,e)\downarrow$

These are, strictly speaking, conditional existence claims, given the free restricted variables, though we know that since concepts exist, unconditional existence claims can be derived from (.1) – (.3).

(**633**) **Definition:** Notation for the Product of Concepts d and e. Given our last theorem, we introduce notation for *the* product of concepts d and e, as follows:

$$d \otimes e =_{df} \imath c ProductOf(c,d,e)$$

As in the case of $\oplus$, expressions of the form $\kappa \otimes \kappa'$ are significant only when it is known, either by proof or by hypothesis, that $C!\kappa$ and $C!\kappa'$.

(**634**) **Lemmas:** Strict Canonicity of Products. Given the definition of *ProductOf*, it is clear that that $d \otimes e$ is (identical to) a canonical concept, for any concepts d and e:

(.1) $d \otimes e = \imath c\forall F(cF \equiv dF \,\&\, eF)$

Where φ is the formula $dF \,\&\, eF$, we can establish the following universal claim by modally strict means and so it tells us that (.2) φ is a rigid condition on properties:

(.2) $\forall F((dF \,\&\, eF) \rightarrow \Box(dF \,\&\, eF))$

Thus $d{\otimes}e$ is a strictly canonical concept, by (260.2). So we can use (261.2) to establish the modally strict theorem that (.3) $d{\otimes}e$ is a concept that encodes just those properties F such that both dF and eF:

(.3) $C!d{\otimes}e \;\&\; \forall F(d{\otimes}eF \equiv dF \,\&\, eF)$

As in the case of $\oplus$, the first conjunct of (.2) is an exemplification formula consisting of a unary relation term $C!$ and a complex individual term $d{\otimes}e$. Finally, it follows that (.4) $d{\otimes}e$ is a product of d and e:

(.4) $\mathit{ProductOf}(d{\otimes}e, d, e)$

(.1) – (.4) are modally strict theorems used in the proofs of subsequent theorems.

(**635**) **Theorems:** Concept Multiplication Forms a Semi-Lattice. It follows immediately from the previous lemma that when we restrict our attention to concepts, $\otimes$ is (.1) idempotent, (.2) commutative, and (.3) associative:

(.1) $c{\otimes}c = c$

(.2) $c{\otimes}d = d{\otimes}c$

(.3) $(c{\otimes}d){\otimes}e = c{\otimes}(d{\otimes}e)$

Thus, concept multiplication, like concept addition, behaves like an algebraic operation.

(**636**) **Theorems:** The Laws of Absorption. We may now prove that laws of *absorption* hold with respect to $\oplus$ and $\otimes$. They are (.1) the sum of c and the product of c and d is identical to c; and (.2) the product of c and the sum of c and d is identical to c:

(.1) $c \oplus (c \otimes d) = c$

(.2) $c \otimes (c \oplus d) = c$

From theorems (621.1) – (621.3), (635.1) – (635.3), and (.1) – (.2) above, we have that (a) $\oplus$ and $\otimes$ both are idempotent, commutative, and associative, and (b) the absorption laws hold. Hence, an algebraist would say that concepts and, hence, abstract objects generally, are structured as a *lattice*, with $\oplus$ as the *join* and $\otimes$ as the *meet* for the lattice.

(**637**) **Theorems:** A Bounded Lattice of Concepts. Indeed, with just a little bit more work, one can show that concepts, and thus abstract individuals generally, are structured as a *bounded* lattice.

First, note that, by the definitions in (263), $Null(x)$ holds whenever x is an abstract object x that encodes no properties, and $Universal(x)$ holds whenever x is an abstract object that encodes every property. We showed that there is exactly one null object and exactly one universal object, and we introduced the notation $\boldsymbol{a}_\varnothing$ (265.1) for $\imath x Null(x)$ and the notation $\boldsymbol{a}_{\boldsymbol{V}}$ (265.2) for $\imath x Universal(x)$. Since $C!$ is defined as $A!$, it follows that there is exactly one null concept and exactly one universal concept. So we may justifiably call $\boldsymbol{a}_\varnothing$ *the null concept* and call $\boldsymbol{a}_{\boldsymbol{V}}$ *the universal concept*.

Consequently the facts proved in (266) allow us to show (.1) the sum of a concept c and the null concept just is c; and (.2) the product of a concept c and the universal concept just is c:

(.1) $c \oplus \boldsymbol{a}_\varnothing = c$

(.2) $c \otimes \boldsymbol{a}_{\boldsymbol{V}} = c$

Thus, $\boldsymbol{a}_\varnothing$ is the identity element for concept addition, and $\boldsymbol{a}_{\boldsymbol{V}}$ is the identity element for concept multiplication:

> **Exercise**. Show $\boldsymbol{a}_\varnothing$ constitutes a minimal element and $\boldsymbol{a}_{\boldsymbol{V}}$ constitutes a maximal element in our lattice of concepts, i.e., show: $\forall c(\boldsymbol{a}_\varnothing \preceq c)$ and $\forall c(c \preceq \boldsymbol{a}_{\boldsymbol{V}})$. [Note: These exercises anticipate some later theorems.]

Consequently, concepts and, hence, abstract objects generally, are structured as a bounded lattice.

Finally, note that (.3) the sum of c and the universal concept is just the universal concept, and (.4) the product of c and the null concept is just the null concept:

(.3) $c \oplus \boldsymbol{a}_{\boldsymbol{V}} = \boldsymbol{a}_{\boldsymbol{V}}$

(.4) $c \otimes \boldsymbol{a}_\varnothing = \boldsymbol{a}_\varnothing$

Thus, no concept survives addition with $\boldsymbol{a}_{\boldsymbol{V}}$ and no concept survives multiplication with $\boldsymbol{a}_\varnothing$.

(**638**) **Remark:** A Boolean Algebra. With just a few more definitions and theorems, we can show that (Leibnizian) concepts with concept addition and multiplication obey the principles of a Boolean algebra. This will occupy our attention over the course of the next few items.

We already know, relative to the domain of concepts, that $\oplus$ and $\otimes$ are idempotent, commutative, and associative, that the absorption laws for $\oplus$ and $\otimes$ hold, that $\boldsymbol{a}_\varnothing$ is an identity element for $\oplus$, and that $\boldsymbol{a}_{\boldsymbol{V}}$ is an identity element for $\otimes$. So it remains only to show that: (i) $\oplus$ distributes over $\otimes$, (ii) $\otimes$ distributes over $\oplus$, and (iii) concept complementation, $-c$, can be defined so that the complementation laws, $c \otimes -c = \boldsymbol{a}_{\varnothing}$ and $c \oplus -c = \boldsymbol{a}_{\boldsymbol{V}}$, both hold. As we shall see, all of these fundamental axioms of Boolean algebra are theorems.

(**639**) **Theorems:** Distribution Laws. Since disjunction distributes over conjunction (88.6.b) and vice versa (88.6.a), it follows that (.1) $\oplus$ distributes over $\otimes$, and (.2) $\otimes$ distributes over $\oplus$:

(.1) $c \oplus (d \otimes e) = (c \oplus d) \otimes (c \oplus e)$

(.2) $c \otimes (d \oplus e) = (c \otimes d) \oplus (c \otimes e)$

(**640**) **Definition:** Complements. We say that c is a *complement of* d whenever c encodes exactly the properties that d fails to encode::

$$ComplementOf(c,d) \equiv_{df} \forall F(cF \equiv \neg dF)$$

For example:

let c_1 be $\imath c \forall F(cF \equiv F = P)$

let c_2 be $\imath c \forall F(cF \equiv F \neq P)$

Since c_1 encodes just the property P and c_2 encodes all and only properties other than P, it follows that $ComplementOf(c_2, c_1)$.

> **Exercise:** Show $ComplementOf(c,d) \to ComplementOf(d,c)$, i.e., that *complement of* is symmetrical.

(**641**) **Theorems:** Facts About Complementation. In the usual way, it follows that (.1) there exists a complement of d; (.2) there exists a unique complement of d; and (.3) the complement of d exists:

(.1) $\exists c ComplementOf(c,d)$

(.2) $\exists! c ComplementOf(c,d)$

(.3) $\imath c ComplementOf(c,d)\downarrow$

These existence claims are conditional on the fact that d is a concept, but since concepts exists, unconditional existence claims are derivable.

(**642**) **Definition:** The Complement of d. Given our last theorem, we are entitled to introduce notation for *the* concept that is a complement of d:

$$-d \;=_{df}\; \imath c ComplementOf(c,d)$$

Since d is a free restricted variable in this definition, an expression of the form $-\kappa$ is significant only when it is known, either by proof or hypothesis, that $C!\kappa$.

(**643**) **Lemmas:** Strict Canonicity of Complements. It is a fact that (.1) $-d$ is identical to a canonical concept:

(.1) $-d = \imath c \forall F(cF \equiv \neg dF)$

Moreover, since the following universal claim is derivable by modally strict means, (.2) metatheoretically implies that $\neg dF$ is a rigid condition on properties:

(.2) $\forall F(\neg dF \rightarrow \Box \neg dF)$

Hence, $-d$ is a strictly canonical concept, by (260.2). So by (261.2), we can, by modally strict means, establish that (.3) $-d$ is a concept that encodes all and only the properties that d fails to encode:

(.3) $C!-d \ \& \ \forall F(-dF \equiv \neg dF)$

Thus, it follows that (.4) $-d$ is a complement of d:

(.4) $ComplementOf(-d, d)$

(.1) – (.4) are modally strict theorems used in the proofs of subsequent theorems.

(**644**) **Theorems:** Complementation Laws. The complementation laws are now theorems. They are (.1) the sum of c and the complement of c is the universal concept, and (.2) the product of c and the complement of c is the null concept:

(.1) $c \oplus -c = \boldsymbol{a}_{\boldsymbol{V}}$

(.2) $c \otimes -c = \boldsymbol{a}_{\varnothing}$

We have now established the commutativity and associativity of $\oplus$ and $\otimes$, the distribution laws for $\oplus$ over $\otimes$ and for $\otimes$ over $\oplus$, the absorption laws, and the complementation laws. Thus, concepts and, hence, abstract objects generally, have the structure of a Boolean algebra.

(**645**) **Theorems:** Other Traditional Principles of Boolean Algebra. We close our discussion of the Boolean algebra of concepts by noting a few final theorems, namely, double complementation and the two De Morgan Laws. These are (.1) the complement of the complement of c just is c; (.2) the sum of $-c$ and $-d$ is identical to the complement of the product of c and d; and (.3) the product of $-c$ and $-d$ is identical to the complement of the sum of c and d:

(.1) $--c = c$

(.2) $-c \oplus -d = -(c \otimes d)$

(.3) $-c \otimes -d = -(c \oplus d)$

Exercises: Show that the following two principles, both of which have been used to axiomatize Boolean algebra, are derivable:

- $-(-x \oplus -y) \oplus -(-x \oplus y) = x$ (Huntington 1933a, Postulate 4.6)
- $-(-(x \oplus y) \oplus -(x \oplus -y)) = x$ (Robbins, reported in McCune 1997)[331]

(**646**) **Exercises:** Concept Difference and Overlap. In 1690a, Leibniz introduces concept subtraction (Theorems VIII – XII). He remarks, in a footnote, that "in the case of concepts, subtraction is one thing, negation another" (1690a, LLP 127). In Theorem IX, he introduces the notion of 'communicating' concepts, by which he seems to mean concepts that in some sense overlap. He then proves the theorem (Theorem X) that if N is the result of subtracting A from L, then A and N are uncommunicating (i.e., don't overlap).

Exercise 1. Define:[332]

(.1) $\mathit{DifferenceOf}(c,d,e) \equiv_{df} \forall F(cF \equiv dF \,\&\, \neg eF)$

(.2) $\mathit{Overlap}(c,d) \equiv_{df} \exists F(cF \,\&\, dF)$

Prove there is a unique concept that is the difference of d and e, that the difference concept of d and e exists, and introduce a restricted term, $d \ominus e$, for that concept:

(.3) $\exists! c\mathit{DifferenceOf}(c,d,e)$

(.4) $\iota c\mathit{DifferenceOf}(c,d,e)\downarrow$

(.5) $d \ominus e =_{df} \iota c\mathit{DifferenceOf}(c,d,e)$

Show that $d \ominus e$ is strictly canonical and prove, as a modally strict theorem, that $d \ominus e$ is a concept that encodes all and only the properties F such that dF and not eF:

(.6) $C!d \ominus e \,\&\, \forall F(d \ominus eF \equiv dF \,\&\, \neg eF)$

Then prove Theorem X in Leibniz 1690a:

(.7) $d \ominus e = c \rightarrow \neg \mathit{Overlap}(e,c)$

---

[331] See also McCune, et al. 2002, who formulate the Robbins axiom slightly differently, namely as $-(-(x \oplus y) \oplus -(-x \oplus y)) = y$.

[332] The notion of *overlap* will play a role in later theorems and it will be officially defined in item (653.1). But it may prove instructive to get some practice in with this notion now.

and show that it is equivalent to:

$$\neg Overlap(e, d\ominus e)$$

Show that the difference of d and e is identical to the product of d and $-e$, i.e., that:

(.8) $d\ominus e = d\otimes -e$

Formulate and prove other theorems governing concept difference and overlap, and determine whether other principles in Leibniz 1690a can be represented and derived. For example, while it is clear how to think about $d \ominus e$ when $e \leq d$, what happens when e is not included in d?

Exercise 2. Consider the scenario in which d and e don't overlap and, in particular, consider any three, pairwise distinct, properties, say P, Q, and R, and the following two objects:

$$d_1 = \imath c\forall F(cF \equiv F = P \vee F = Q)$$

$$e_1 = \imath c\forall F(cF \equiv F = R)$$

Give a systematic answer to the questions, what properties does $d_1 \ominus e_1$ encode, and what properties does $e_1 \ominus d_1$ encode?

Exercise 3. Consider the scenario in which d is included in e but not identical to e and, in particular, consider:

$$d_2 = \imath c\forall F(cF \equiv F = P)$$

$$e_2 = \imath c\forall F(cF \equiv F = P \vee F = Q)$$

Identify $d_2 \ominus e_2$ and $e_2 \ominus d_2$ in terms of particular concepts we've already discussed.

Exercise 4. Finally, prove:

(.9) $d = e \equiv (d \ominus e = e \ominus d)$

(.10) $Overlap(c, d) \equiv c \otimes d \neq \boldsymbol{a}_\varnothing$

(.11) $\neg Overlap(c, d) \equiv c \ominus d = c$

13.1.5 The Mereology of Concepts

(**647**) **Remark:** Mereology. Now that we have established that concepts form a Boolean algebra, we consider the ways in which they constitute a mereology, i.e., the ways in which the notions of *part* and *whole* can be defined and applied to concepts. Though some authors take mereology to apply primarily to the domain of concrete individuals (e.g., Simons 1987, 4),[333] others suggest that mereology makes no assumptions about the kinds of entities having parts. Thus, Varzi (2015, §1) writes:

> ... it is worth stressing that mereology assumes no ontological restriction on the field of 'part'. In principle, the relata can be as different as material bodies, events, geometric entities, or spatio-temporal regions, ... as well as abstract entities such as properties, propositions, types, or kinds, ... As a formal theory ... mereology is simply an attempt to lay down the general principles underlying the relationships between an entity and its constituent parts, whatever the nature of the entity, just as set theory is an attempt to lay down the principles underlying the relationships between a set and its members. Unlike set theory, mereology is not committed to the existence of abstracta: the whole can be as concrete as the parts. But mereology carries no nominalistic commitment to concreta either: the parts can be as abstract as the whole.

Moreover, it is often suggested that the entities of a mereological domain are structured algebraically. Simons notes that "the algebraic structure of a full classical mereology is that of a complete Boolean algebra with zero deleted" (1987, 25).[334] Consequently, since (a) mereological principles can be understood broadly as applying to abstract objects such as concepts, (b) concepts are abstract objects, and (c) the previous section shows that the domain of concepts is structured algebraically, it seems reasonable to investigate the extent to which concepts are provably governed by mereological principles.

Our discussion in the remainder of this section will be organized as follows. We begin by examining what happens when we interpret the mereological notion *part of* as the *inclusion* ($\preceq$) condition on concepts. That is, we confirm that core mereological principles are preserved when we both (a) define c is a *part of* d just in case d encodes every property c encodes and (b) define *proper part of* in the usual mereological way. Then we examine a variety of consequences of our definitions and consider whether they are acceptable as mereological principles. Finally, we examine what mereological principles provably apply to *non-null* concepts, i.e., concepts that encode at least one property. We discover that while some questionable mereological principles apply to concepts generally, they do not apply to non-null concepts.

In what follows, we assume familiarity with the basic notions and principles of mereology. Systems of mereology are typically, though not always, formulated in one of two ways. Some systems take *x is a part of y* ($x \preceq y$) as

[333] However, in 1987 (169–171), Simons does offer a few brief thoughts about mereology and abstract objects.

[334] A zero element is defined mereologically as an individual that is a part of every individual. It is traditional to suppose that there is no such zero element in a domain of concrete individuals.

a primitive relation or as a relational condition that is, at a minimum, reflexive, anti-symmetric, and transitive. Others start with *x is a proper part of y* ($x \prec y$) as the primitive relation (or relational condition) that is, at a minimum, irreflexive, asymmetric, and transitive.[335] Little hangs on the choice of formulation since it is a well-known fact that to every non-strict partial order (e.g., one based on $\preceq$) that is reflexive, anti-symmetric and transitive, there corresponds a strict partial order (e.g., one based on $\prec$) that is provably irreflexive, asymmetric, and transitive. To obtain $\prec$ when starting with $\preceq$, one defines $x \prec y$ as $x \preceq y \,\&\, x \neq y$. Then the irreflexivity, asymmetry, and transitivity of $\prec$ follow from facts about $\preceq$ and $\neq$. Alternatively, to obtain $\preceq$ when starting with $\prec$, one defines $x \preceq y$ as $x \prec y \vee x = y$. Then the reflexivity, anti-symmetry and transitivity of $\preceq$ follow from facts about $\prec$ and $=$. We begin by seeing how these ideas are confirmed in the present theory, under object-theoretic definitions of the notions involved.

(**648**) **Remark:** Part Of. Though mereology often takes *part of* as a primitive, we may define it object-theoretically and show that its core features are derivable. So let us *re*-introduce definition (624.1) but in such a way that the definiendum, $c \preceq d$, is to be read: concept c is a *part of* concept d. Thus, the following definition now systematizes a different pretheoretical notion:

$$c \preceq d \ \equiv_{df} \ \forall F(cF \rightarrow dF) \qquad (624.1)$$

Note that our system does not guarantee that $\preceq$ defines a relation on concepts.[336] Nevertheless, $c \preceq d$ is a well-defined binary condition on concepts and, indeed, on abstract objects generally.

Given the above definition, the theorems in (625) guarantee that *part of* ($\preceq$) is reflexive, anti-symmetric and transitive with respect to the concepts:

(.1) $c \preceq c$ (625.1)
A concept is a part of itself.

(.2) $c \preceq d \rightarrow (c \neq d \rightarrow d \npreceq c)$ (625.2)
If a concept c is a part of a distinct concept d, then d is not a part of c.

[335] There are other ways of formulating a mereology, for example, by taking *x overlaps y* as a primitive (Goodman 1951), or by taking *x is disjoint from y* as a primitive (Leonard and Goodman 1940). But these variations need not distract us in what follows. See Simons 1987, p. 48ff.

[336] At least, axiom (39.2) doesn't guarantee that there is such a relation. The definition uses restricted variables, and so is shorthand for:

$$x \preceq y \ \equiv_{df} \ C!x \,\&\, C!y \,\&\, \forall F(xF \rightarrow yF)$$

By the Convention for Encoding Formulas (17.3), the variables x and y occur free in encoding position in $x \preceq y$ since they occur free in encoding position in the definens. So $[\lambda xy\, x \preceq y]$ is not a core λ expression, since the λ binds variables that occur in encoding position in the matrix. So by our convention for restricted variables, $[\lambda cd\, c \preceq d]$ also fails to be a core λ-expression. Neither λ-expression is asserted to be significant by (39.2).

(.3) $c \preceq d \;\&\; d \preceq e \;\rightarrow\; c \preceq e$ (625.3)
If c is part of d and d is part of e, then c is part of e.

Thus, it is clear that the above definition of $\preceq$ captures the notion of *improper part of*.

Varzi (2015) notes that principles such as these "represent a common starting point of all standard [mereological] theories" (2015, §2.2). However, he notes later that "[n]ot just any partial ordering qualifies as a part-whole relation, though, and establishing what further principles should be added ... is precisely the question a good mereological theory is meant to answer" (2015, §3). We shall return to this question below.

(**649**) **Definitions:** Proper Part Of. We define *proper part of* in the usual way, though relative to the domain of concepts. We say concept c is a *proper part* of concept d, written $c \prec d$, just in case c is a part of d and c is not equal to d:

$$c \prec d \;\equiv_{df}\; c \preceq d \;\&\; c \neq d$$

Warning! In what follows, we sometimes cite Simons' classic text of 1987, which uses $<$ and $\ll$ for improper and proper parthood, respectively. Thus, our symbol $\preceq$ for improper parthood corresponds to his symbol $<$, and our symbol $\prec$ for proper parthood corresponds to his symbol $\ll$. So it is important not to confuse our symbol for proper parthood ($\prec$) with his symbol for improper parthood ($<$).

(**650**) **Theorems:** Principles of Proper Parthood. As expected, we can now derive that *proper parthood* is a strict partial ordering with respect to concepts. It immediately follows that (.1) c is not a proper part of itself; (.2) if c is a proper of d and d is a proper part of e, then c is a proper part of e; and (.3) if c is a proper part of d, then d is not a proper part of c:

(.1) $c \not\prec c$ (Irreflexivity)

(.2) $c \prec d \;\&\; d \prec e \;\rightarrow\; c \prec e$ (Transitivity)

(.3) $c \prec d \rightarrow d \not\prec c$ (Asymmetry)

Simons notes that these principles "fall well short of characterizing the [proper] part-relation; there are many [strict] partial orderings which we should never call part-whole systems" (1987, 26). In what follows, we examine the extent to which our theorems conform to the accepted principles of part-whole systems, by considering the broader picture of how *traditional* mereological notions fare in the current setting.

(**651**) **Definitions:** Bottom Element, Concepts With a Single Proper Part, and Atoms. Three of the most basic mereological issues are: (a) whether there exists an individual that is a part of every individual (i.e., whether there exists a 'bottom' or 'zero' element), (b) whether there are any individuals that have a single proper part (i.e., exactly one proper part), and (c) whether there exist any individuals that have no proper parts (i.e., whether there exist any 'atoms').

The traditional mereology of concrete individuals eschews the existence of bottom elements and individuals with exactly one proper part, though the existence of atoms is permitted. For example, Simons writes (1987, 13):

> ... in normal set theory even two disjoint sets have an intersection, namely the null set, whereas disjoint individuals precisely lack any common part. Most mereological theories have no truck with the fiction of a null individual which is part of all individuals, although it neatens up the algebra somewhat.

Clearly, the preference Simons is describing seems reasonable when one supposes, as he does, that mereology is restricted to the study of the part-whole relation on concrete individuals. But that is not the case in the present context.

Similarly, traditional mereology eschews concrete wholes having a single proper part. Simons rhetorically asks and then asserts, "How could an individual have a *single* proper part? That goes against what we mean by 'part'" (1987, 26). Again, this view seems reasonable when mereology is limited to the field of concrete objects. But it is worth noting that Varzi (2015, §3.1) compiles a list of objects with a single proper part that have been postulated by philosophers.

By contrast, in our algebra of abstract individuals, it follows that there is a unique bottom concept and that there are concepts having a single proper part. Moreover, there are conceptual atoms. To see that these are facts, let us begin by saying that (.1) c is a *bottom* concept just in case c is a part of every concept; and (.2) c is an *atom* just in case c has no concepts as proper parts:

(.1) $Bottom(c) \equiv_{df} \forall d(c \preceq d)$

(.2) $Atom(c) \equiv_{df} \neg\exists d(d \prec c)$

In these definitions, then, standard mereological notions have been adapted to the present context (Simons 1987, 16; Varzi 2015, §3.4).

(**652**) **Theorems:** Facts About Bottom Concepts, Concepts With a Single Proper Part, and Atoms. Recall that we relabeled $\boldsymbol{a}_{\varnothing}$ as *the null concept* (637). Since modally-strict theorem (266.3), i.e., $Null(\boldsymbol{a}_{\varnothing})$, implies that the null concept encodes no properties, we can now prove: (.1) the null concept is a bottom concept; (.2) there is a unique bottom concept; (.3) the null concept is a proper part of the Thin Form of G ($\boldsymbol{a}_G$); and (.4) the thin Form of G has exactly one proper part; (.5) bottom concepts are atoms; (.6) the null concept is an atom; and (.7) there is exactly one atom:

(.1) $Bottom(\boldsymbol{a}_\varnothing)$

(.2) $\exists! cBottom(c)$

(.3) $\boldsymbol{a}_\varnothing \prec \boldsymbol{a}_G$

(.4) $\exists! d(d \prec \boldsymbol{a}_G)$

(.5) $Bottom(c) \rightarrow Atom(c)$

(.6) $Atom(\boldsymbol{a}_\varnothing)$

(.7) $\exists! cAtom(c)$

Later, we shall return to the discussion of bottom concepts, concepts with exactly one proper part, and atoms; if we study how *part of* behaves when restricted to *non-null concepts* (i.e., concepts that encode at least one property), then there are related notions of bottoms and atoms that do not yield the above consequences. To give a hint of what is to come, $\boldsymbol{a}_\varnothing$ fails to be a *non-null bottom* (i.e., a non-null concept that is a part of every non-null concept) since it fails to be a non-null concept; indeed, there is no non-null bottom element. Moreover, $\boldsymbol{a}_\varnothing$ fails to be a *non-null atom* (i.e., a non-null concept that has no proper parts), though as we shall see, there nevertheless are non-null atoms. But we shall discuss these ideas in more detail below, starting with item (659).

(**653**) **Definitions:** Mereological Overlap. In Exercise 1 of (646), we considered the following definition: *c overlaps d* whenever there is a property F that both c and d encode:

(.1) $Overlap(c,d) \equiv_{df} \exists F(cF \,\&\, dF)$

Clearly, this is a natural understanding of overlap in our object-theoretic setting. However, readers familiar with texts on mereology will recognize that the standard definition of *mereological overlap* is different from (.1). The standard mereological definition is that individuals x and y overlap whenever they have a common part (Simons 1987, 11, 28; Varzi 2015, §2.2). Let us formulate this latter definition in the present context by saying that *c overlaps* d* just in case there is a concept e that is a part of both concepts c and d:

(.2) $Overlap^*(c,d) \equiv_{df} \exists e(e \preceq c \,\&\, e \preceq d)$

These two notions of overlap have some interesting consequences.

(**654**) **Theorems:** Overlap vs. Overlap*. We first observe that the notions of overlap and overlap* are not equivalent; (.1) overlap implies overlap*, but (.2) overlap* doesn't imply overlap:

(.1) $Overlap(c,d) \rightarrow Overlap^*(c,d)$

(.2) $\exists c \exists d(Overlap^*(c,d) \,\&\, \neg Overlap(c,d))$

Clearly, since the null concept $\boldsymbol{a}_\varnothing$ encodes no properties, it overlaps with nothing and so doesn't overlap with itself. Hence it follows that (.3) overlap is not reflexive:

(.3) $\neg \forall c Overlap(c,c)$

But, trivially, since every concept is a part of itself, (.4) overlap* is reflexive:

(.4) $Overlap^*(c,c)$

By contrast, both overlap and overlap* are symmetric:

(.5) $Overlap(c,d) \rightarrow Overlap(d,c)$

(.6) $Overlap^*(c,d) \rightarrow Overlap^*(d,c)$

More conclusively, though, whereas (.7) not all concepts overlap each other, (.8) all concepts overlap* each other:

(.7) $\neg \forall c \forall d Overlap(c,d)$

(.8) $\forall c \forall d Overlap^*(c,d)$

(.8) is a consequence of the definition of overlap* and the fact that the null concept, $\boldsymbol{a}_\varnothing$, is a bottom element: since $\boldsymbol{a}_\varnothing$ is a part of every concept, it is a common part of every two concepts. Given the mereological commitment to the definition of *overlap* as *having a common part*, a consequence such as (.8) makes it clear why traditional mereologists eschew a *bottom* element. Varzi thus notes that "[I]n general ... mereologists tend to side with traditional wisdom and steer clear of (P.10) [which asserts the existence of a bottom element] altogether" (2015, §3.4). We shall see, however, that mereological overlap again becomes a useful condition when we turn our attention to the non-null concepts – mereological overlap doesn't hold universally on that subdomain.

Let's return, then, to the object-theoretic definition of overlap, whereby concepts overlap if they encode a common property. Then note that for arbitrary concepts c and d, (.9) there is a concept e that is a part of d but which doesn't overlap with c:

(.9) $\exists e(e \preceq d \,\&\, \neg Overlap(e,c))$

(The reason for ordering the free variables in this way will become apparent when we discuss supplementation principles below.) While it is straightforward to derive (.9) when $\boldsymbol{a}_\varnothing$ is taken as a witness, the theorem also derivable when $d \ominus c$ is selected as the witness; see the proof in the Appendix. Furthermore, the theorem holds even in the degenerate case where the variables c and

d take $\boldsymbol{a}_\varnothing$ as their value; in this case, the witness is still $\boldsymbol{a}_\varnothing$ (which is then identical to $d \ominus c$), since $\boldsymbol{a}_\varnothing$ is a part of itself and fails to overlap with itself.

Finally, note that the principle which results when we replace $\preceq$ in (.9) with $\prec$ doesn't hold for arbitrary concepts c and d. We have:

(.10) $\neg\forall c \forall d \exists e(e \prec d \,\&\, \neg Overlap(e,c))$

I.e., $\exists c \exists d \neg\exists e(e \prec d \,\&\, \neg Overlap(e,c))$. To see this, consider any concept c as a witness to the first existential quantifier and consider $\boldsymbol{a}_\varnothing$ as a witness to the second. Then since $\boldsymbol{a}_\varnothing$ is an atom, it has no proper parts, i.e., $\neg\exists e(e \prec \boldsymbol{a}_\varnothing)$. *A fortiori*, $\neg\exists e(e \prec \boldsymbol{a}_\varnothing \,\&\, \neg Overlap(e,c))$.

(**655**) **Theorems:** Supplementation Principles. Both weak and strong supplementation principles of mereology trivially follow from (654.9). The *weak supplementation* principle of mereology is that if an individual x is a proper part of individual y, then there exists an individual that is both a part of y and fails to overlap x (Varzi 2015, §3.1, item P.4). The *strong supplementation* principle is that if y fails to be a part of x, then there is an individual that is a part of y and that fails to overlap x (Varzi 2015, §3.2, item P.5; Simons 1987, 29, SA5/SSP). When we consider the domain of concepts, these become the theorems that (.1) if c is a proper part of d, then there is a concept e such that e is a part of d and e fails to overlap c, and (.2) if d fails to be a part of c, then there is a concept e such that e is a part of d and fails to overlap c:

(.1) $c \prec d \rightarrow \exists e(e \preceq d \,\&\, \neg Overlap(e,c))$

(.2) $d \npreceq c \rightarrow \exists e(e \preceq d \,\&\, \neg Overlap(e,c))$

Since the consequents of both of these conditionals are just instances of (654.9), both theorems follow trivially. Given that (.1) and (.2) are theorems independent of the truth of their antecedents, a mereologist will no doubt question whether *overlap*, as defined in (653.1), is a proper mereological notion. This is a fair question, though it will become clear that when restricted to the subdomain of non-null concepts, this notion of *overlap* is a proper mereological notion and, indeed, becomes equivalent to the standard mereological notion.

However, note that the following variant of weak supplementation is a nontrivial theorem, namely, that (.3) if c is a proper part of d, then there is a concept e that is a proper part of d that fails to overlap with c:

(.3) $c \prec d \rightarrow \exists e(e \prec d \,\&\, \neg Overlap(e,c))$

We saw in (654.10) that the consequent of the above claim doesn't hold for arbitrary concepts c and d. So (.3) isn't a theorem merely in virtue of the truth of the consequent. Note, however, that one can't take the witness to the consequent to be $d \ominus c$. For in the case where c is the null concept $\boldsymbol{a}_\varnothing$, it is not the case that $d \ominus c \prec d$, since $d \ominus c$ just is d. So the proof of (.3) requires that the witness to the consequent be $\boldsymbol{a}_{\varnothing}$. Hence, one can argue that (.3) is not a true supplementation principle, since it is reasonable to suppose that $\boldsymbol{a}_{\varnothing}$ doesn't supplement anything. As we shall see, however, when we restrict the variables in (.3) to non-null concepts, the resulting principle (668.3) does have claim to being a supplementation principle.

(**656**) **Definition:** Underlap and Maximal Concepts. The object-theoretic notion of underlap is that c and d underlap just in case some concept encodes all the properties c encodes as well as all the properties d encodes. By contrast, in traditional mereology, two individuals underlap just in case there is an individual of which they are both a part. But we won't formalize and distinguish both notions of underlap since it turns out they are equivalent; the definiens of both definitions are theorems.[337] Instead, we shall just work with the traditional, mereological notion of underlap, namely (.1) c and d *underlap* just in case there is a concept e such that both c and d are a part of e:

(.1) $Underlap(c,d) \equiv_{df} \exists e(c \leq e \,\&\, d \leq e)$

Moreover, let us define (.2) a *maximal concept* to be any concept of which every concept is a part:

(.2) $MaxConcept(c) \equiv_{df} \forall d(d \leq c)$

(**657**) **Theorem:** Maximal Concepts and Underlap. Clearly, it is provable that (.1) the universal object, $\boldsymbol{a}_{\boldsymbol{V}}$, is a maximal concept:

(.1) $MaxConcept(\boldsymbol{a}_{\boldsymbol{V}})$

It is a well-known fact of mereology that if there exists a universal whole, of which every individual is a part, then every individual underlaps every individual.[338] But even without a maximal concept, one can cite the sum $c \oplus d$ as a witness that verifies that (.2) underlap holds universally among concepts:

[337] The two definitions may be formalized as follows:

$Underlap_1(c,d) \equiv_{df} \exists e \forall F((cF \vee dF) \rightarrow eF)$ (object-theoretic notion)

$Underlap_2(c,d) \equiv_{df} \exists e(c \leq e \,\&\, d \leq e)$ (mereological notion)

Then it is easy to show: $Underlap_1(c,d) \equiv Underlap_2(c,d)$. Both the 'left' and 'right' conditional hold by virtue of the truth of the consequent. ($\rightarrow$) To show $Underlap_1(c,d) \rightarrow Underlap_2(c,d)$, it suffices to show $Underlap_2(c,d)$. For the latter, it suffices to show that $c \oplus d$ is a witness to $\exists e(c \leq e \,\&\, d \leq e)$. But clearly, by (627.1) and (627.2), respectively, we know both $c \leq c \oplus d$ and $d \leq c \oplus d$. ($\leftarrow$) To show $Underlap_2(c,d) \rightarrow Underlap_1(c,d)$, it suffices to show $Underlap_1(c,d)$. For the latter, it suffices to show that $c \oplus d$ is a witness to $\exists e \forall F((cF \vee dF) \rightarrow eF)$. So, by GEN, we need to show $(cF \vee dF) \rightarrow c \oplus dF$. But this follows *a fortiori* from the second conjunct of (620.3), which in the present case implies $c \oplus dF \equiv cF \vee dF$.

[338] Cf. Varzi 2015, §2.2, where it is noted that "Uxy is bound to hold if one assumes the existence of a 'universal entity' of which everything is part."

(.2) $\forall c \forall d Underlap(c, d)$

A proof of this claim appeared in footnote 337, where the equivalence of two concepts of underlap was discussed.

(**658**) **Remark:** Towards an Uncontroversial Mereology. We've now seen a variety of new theorems, with $\leq$ considered as *improper parthood*. Some of those theorems, so understood, preserve key mereological principles, while others may be thought doubtful as such. The main source of doubt arises from (the consequences of) the facts that the null concept $\boldsymbol{a}_{\varnothing}$ is a bottom concept that is a part of every concept and that the definition of overlap, which is stated in terms of encoding rather than in terms of the mereological notion of parthood, yields unexciting supplementation principles.

But as mentioned previously, one can sidestep these concerns by excluding the null concept from the algebra, so that it is no longer in the domain of the mereology. There are at least two ways in which this can be done. Varzi suggests (2015, §3.4) that one "treat the null item as a mere algebraic fiction" and revise the definition of parthood as follows:

Genuine Parthood
$GP(x, y) \equiv_{df} x \leq y \,\&\, \exists z \neg (x \leq z)$

To understand the suggestion better, note that the second conjunct of the definiens is equivalent to $\neg \forall z (x \leq z)$. Given that $Bottom(x, z)$ is defined as $\forall z (x \leq z)$, the above definition becomes: x is a *genuine part* of y iff x is a part of y and not a bottom element. Now in the present theory of concepts, we know that there is a unique bottom concept, namely, $\boldsymbol{a}_{\varnothing}$. So in the context of our theory, Varzi's suggestion could be implemented by defining: c is a *genuine part* of d just in case c is a part of d other than $\boldsymbol{a}_{\varnothing}$ (or just in case c is part of d and c encodes some property). Varzi then suggests a definition of 'genuine overlap' in terms of genuine parthood; in object theory, his definition would be implemented as follows: *c genuinely overlaps d* if and only if there is a concept *e* that is a genuine part of both.

While this is a perfectly legitimate course to follow, in the present theory, it makes better sense to hold the notions of parthood and proper parthood fixed and consider how they behave on a subdomain, i.e., approach the matter by investigating how the defined notions of overlap, bottom, atom, etc., behave when restricted to the concepts that encode at least one property, i.e., with respect to *non-null* concepts. We now turn to a development of this idea.

(**659**) **Definition:** Non-null Concepts. We may use the definition of a null object ($Null(x)$) in (263.1) to stipulate that x is a *non-null concept* ('$Concept^{+}(x)$'), just in case x is a concept and x is not a null object:

$$Concept^{+}(x) \equiv_{df} C!x \,\&\, \neg Null(x)$$

Though we could just as well have said that a non-null concept is any concept other than the null concept $\boldsymbol{a}_\varnothing$, the above definition avoids the additional layer of complexity that would be contributed by using negation, identity, and $\boldsymbol{a}_\varnothing$ in the second conjunct of the definiens.

(**660**) **Theorem:** $Concept^+(x)$ is a Rigid Restriction Condition. $Concept^+(x)$ is a restriction condition, as defined in (336): it has a single free variable, it is a theorem that $\exists x Concept^+(x)$, and $Concept^+(\kappa) \rightarrow \kappa\downarrow$. But $Concept^+(x)$ is also a rigid restriction condition, since it is a modally strict theorem that, for any object x, if $Concept^+(x)$ then necessarily $Concept^+(x)$:

$$\forall x(Concept^+(x) \rightarrow \Box Concept^+(x)$$

Since it is also clear that concepts$^+$ are concepts, we may introduce rigid (doubly) restricted variables $\underline{c}, \underline{d}, \underline{e}, \ldots$ to range over concepts$^+$. These are doubly restricted because we may eliminate the variables in one of two ways, as discussed in (514). For example, we may regard $\forall \underline{c}\varphi^{\underline{c}}_x$ either as $\forall x(Concept^+(x) \rightarrow \varphi)$ or as $\forall c(Concept^+(c) \rightarrow \varphi^c_x)$.

(**661**) **Definitions:** Non-null Bottoms. Let us say that a non-null concept $\underline{c}$ is a *non-null bottom*, written $Bottom^+(\underline{c})$, just in case $\underline{c}$ is a part of every non-null concept:

$$Bottom^+(\underline{c}) \equiv_{df} \forall \underline{d}(\underline{c} \preceq \underline{d})$$

(**662**) **Theorems:** Facts About Non-null Bottoms. The foregoing definition has the following consequences: (.1) The null concept is not a non-null bottom; and (.2) no non-null concept is a non-null bottom:

(.1) $\neg Bottom^+(\boldsymbol{a}_\varnothing)$

(.2) $\neg\exists \underline{c} Bottom^+(\underline{c})$

So the mereologist's expectation that there be no bottom element is met when we restrict our attention to non-null concepts (cf. Simons 1987, 13, 25; Varzi 2015, §3.4).

(**663**) **Definition:** Non-null Atoms. Let us say that a non-null concept $\underline{c}$ is a *non-null atom*, written $Atom^+(\underline{c})$, just in case $\underline{c}$ has no non-null concepts as proper parts:

$$Atom^+(\underline{c}) \equiv_{df} \neg\exists \underline{d}(\underline{d} \prec \underline{c})$$

(**664**) **Theorems:** Facts About Non-null Atoms and Proper Parthood for Non-null Concepts. We now have the following consequences: (.1) the Thin Form of G is a non-null atom; (.2) no non-null concept is a unique proper part of a non-null concept; (.3) non-null atoms encode at most one property; and (.4) a non-null concept is a non-null atom if and only if it is a Thin Form:

(.1) $Atom^{+}(\boldsymbol{a}_G)$

(.2) $\neg\exists!\underline{d}(\underline{d} \prec \underline{c})$

(.3) $Atom^{+}(\underline{c}) \rightarrow \forall F \forall G((\underline{c}F \;\&\; \underline{c}G) \rightarrow F = G)$

(.4) $Atom^{+}(\underline{c}) \equiv ThinForm(\underline{c})$

Note that (.2) preserves Simon's (1987, 26) intuition, mentioned earlier, that no mereological individual has a single proper part.

(**665**) **Theorem:** Overlap on Non-null Concepts Is Mereological. By restricting our attention to non-null concepts, it emerges that non-null concepts $\underline{c}$ and $\underline{d}$ overlap just in case there is a non-null concept $\underline{e}$ that is a part of both:

$$Overlap(\underline{c}, \underline{d}) \equiv \exists \underline{e}(\underline{e} \preceq \underline{c} \;\&\; \underline{e} \preceq \underline{d})$$

Thus, the mereological definition of *overlap* becomes derivable as a theorem of the subtheory of non-null concepts. This merits a brief discussion.

(**666**) **Remark:** Object-Theoretic and Mereological Overlap on Non-null Concepts. The previous theorem shows that, when we restrict our attention to non-null concepts, the object-theoretic definition of *overlap* defined in (653.1) is equivalent to the traditional mereological definition of overlap defined in (653.2) as *overlap**. Consequently, as we complete our study of non-null concepts, we may regard object-theoretic overlap as a *bona fide* mereological notion.

Of course, lots of questions now arise. For example: (i) Which of the questionable mereological theorems about overlapping concepts can be preserved as *bona fide* mereological principles about overlapping non-null concepts? (ii) What mereological principles of supplementation hold with respect to non-null concepts? These and other questions will be investigated below.

(**667**) **Theorems:** Facts About Overlap and Non-null Concepts. (.1) Some non-null concepts overlap and some don't; (.2) overlap is reflexive on the non-null concepts; (.3) overlap is symmetric on the non-null concepts; (.4) overlap is not transitive on the non-null concepts; (.5) it is not the case that for any non-null concepts c and d, there is non-null concept e that is a part of d but which doesn't overlap c; and (.6) it is not the case that for any non-null concepts c and d, there is non-null concept e that is a proper part of d but which doesn't overlap c:

(.1) $\exists \underline{c} \exists \underline{d} Overlap(\underline{c}, \underline{d}) \;\&\; \exists \underline{c} \exists \underline{d} \neg Overlap(\underline{c}, \underline{d})$

(.2) $Overlap(\underline{c}, \underline{c})$

(.3) $Overlap(\underline{c}, \underline{d}) \rightarrow Overlap(\underline{d}, \underline{c})$

(.4) $\exists\underline{c}\exists\underline{d}\exists\underline{e}(Overlap(\underline{c},\underline{d})\ \&\ Overlap(\underline{d},\underline{e})\ \&\ \neg Overlap(\underline{c},\underline{e}))$

(.5) $\neg\forall\underline{c}\forall\underline{d}\exists\underline{e}(\underline{e}\preceq\underline{d}\ \&\ \neg Overlap(\underline{e},\underline{c}))$

(.6) $\neg\forall\underline{c}\forall\underline{d}\exists\underline{e}(\underline{e}\prec\underline{d}\ \&\ \neg Overlap(\underline{e},\underline{c}))$

Contrast (.1) with (654.7) and (654.8). If we consider concepts generally, then the object-theoretic overlap doesn't hold between every pair of concepts (654.7), while the mereological overlap does (654.8). But since object-theoretic overlap is equivalent to mereological overlap with respect to non-null concepts (665), (.1) reminds us that object-theoretic overlap is not trivial with respect to the non-null concepts.

(.5) is of interest because it contrasts with (654.9). The latter implies that for any two concepts c and d, there is a concept e that is a part of d but which doesn't overlap with c. But (.5) asserts that this fails to hold for non-null concepts; there are non-null concepts $\underline{c}$ and $\underline{d}$ such that every non-null part of $\underline{d}$ mereologically overlaps with $\underline{c}$. Since a principle like (654.9) doesn't hold for arbitrary non-null concepts $\underline{c}$ and $\underline{d}$, the way is now clear to prove *non-trivial* supplementation principles.

Note that (.6) shows that (654.10) remains a theorem when restricted to non-null concepts, albeit for a different reason. To see why (.6) holds, consider its equivalent form: $\exists\underline{c}\exists\underline{d}\forall\underline{e}(\underline{e}\prec\underline{d}\rightarrow Overlap(\underline{e},\underline{c}))$. Then as a witness to the first existential quantifier, take any non-null concept you please, say $\underline{c}_1$. And as a witness to the second existential quantifier, pick an arbitrary property, say G, and consider the Thin Form of G, i.e., $\boldsymbol{a}_G$, which is clearly a non-null concept. Then, by GEN, it suffices to show $\underline{e}\prec\boldsymbol{a}_G\rightarrow Overlap(\underline{e},\underline{c}_1)$. But this is provable by failure of the antecedent: since $\boldsymbol{a}_G$ is a non-null atom (664.1), it has no non-null concepts as proper parts, i.e., $\neg\underline{e}\prec\boldsymbol{a}_G$, by definition (663).

(**668**) **Theorems:** Non-trivial Supplementation Principles on Non-null Concepts. The following weak and strong supplementation principles, formulated so that they apply only to non-null concepts, are not trivial: (.1) if $\underline{c}$ is a proper part of $\underline{d}$, then some non-null part of $\underline{d}$ fails to overlap $\underline{c}$; (.2) if $\underline{d}$ fails to be a part of $\underline{c}$, then some non-null part of $\underline{d}$ fails to overlap $\underline{c}$:

(.1) $\underline{c}\prec\underline{d}\rightarrow\exists\underline{e}(\underline{e}\preceq\underline{d}\ \&\ \neg Overlap(\underline{e},\underline{c}))$ (Weak Supplementation)

(.2) $\underline{d}\npreceq\underline{c}\rightarrow\exists\underline{e}(\underline{e}\preceq\underline{d}\ \&\ \neg Overlap(\underline{e},\underline{c}))$ (Strong Supplementation)

Cf. Varzi 2015, §3.1, P.4, and §3.2, P.5. Finally, note that we obtain the following when we restrict (655.3) to non-null concepts:

(.3) $\underline{c}\prec\underline{d}\rightarrow\exists\underline{e}(\underline{e}\prec\underline{d}\ \&\ \neg Overlap(\underline{e},\underline{c}))$

This is slightly stronger than (.1). Varzi (2015, §3.1, P.4′) calls (.3) *proper supplementation*. It is immune to the concern raised about (655.3); it is a true supplementation principle; the witness to the quantifier in the consequent is not the null concept, as in the proof of (655.3).

(**669**) **Exercises:** Consider the following questions (ranging from easy to hard), all of which arise by considering how previous definitions fare when (adjusted and) applied to non-null concepts:

- Is there a maximal non-null concept, i.e., a $\underline{c}$ such that $\forall \underline{d}(\underline{d} \preceq \underline{c})$?

- Consider the following mereological definition of concept summation (Simons 1987, 32, cf. SD7; Varzi 2015, §4.2, 39_3):

$$Sum^{+}(\underline{c}, \underline{d}, \underline{e}) \equiv_{df} \forall \underline{f}(Overlap(\underline{c}, \underline{f}) \equiv (Overlap(\underline{f}, \underline{d}) \vee Overlap(\underline{f}, \underline{e}))$$

  How does this mereological notion of concept summation relate to the object-theoretic notion we defined in (617) when the latter is restricted to non-null concepts? Are the two notions equivalent? If not, what distinctive theorems do the non-equivalent notions give rise to? For example, we established that *every* two concepts have a unique sum (618.2); does this hold for non-null concepts on the alternative definition of concept summation?

- What lattice-theoretic and algebraic facts are preserved when we restrict our attention to non-null concepts? What facts of this kind fail to hold with respect to non-null concepts? Are the non-null concepts structured in the way Simons says a 'full classical mereology' is structured, namely, as a "complete Boolean algebra with zero deleted" (1987, 25)?

- Finally, does the following *complementation* principle hold:

$$\underline{c} \npreceq \underline{d} \rightarrow \exists \underline{e} \forall \underline{f}(\underline{f} \preceq \underline{e} \equiv (\underline{f} \preceq \underline{c} \,\&\, \neg Overlap(\underline{f}, \underline{d})))$$

  (cf. Varzi 2015, §3.3, P.6)?

We leave the above as open questions for the interested reader to pursue.

13.2 Concepts of Properties and Individuals

In the subsections that follow, we describe: (a) concepts of properties, such as the concept of *being a king*, the concept of *being red*, etc., (b) concepts of ordinary individuals, such as the concept of Adam, the concept of Alexander, etc., and (c) generalized concepts, such as the concept of *every human*, the concept of *something red*, etc.

13.2.1 Concepts of Properties

(**670**) **Definitions:** Concepts of Properties. In (442) we defined, *G necessarily implies F*, written $G \Rightarrow F$, just in case $\Box\forall x(Gx \rightarrow Fx)$. Let us now say that c is a *concept of G* just in case c encodes exactly the properties necessarily implied by G:

$$ConceptOf(c,G) \equiv_{df} G\downarrow \,\&\, \forall F(cF \equiv G \Rightarrow F)$$

We assume our conventions for free restricted variables and for definitions by equivalence.[339] The above definition makes it clear that, for any property G, any concept c such that $ConceptOf(c,G)$ is an individual.

To develop a familiar example, suppose we have extended object theory by adding both the primitive property *being human* and some non-logical axioms that tell us what properties *being human* necessarily implies. For example, suppose we've added the axiom: *being human* necessarily implies *being concrete*, i.e., $H \Rightarrow E!$. Then the definition tells us that $ConceptOf(c,H)$ just in case $\forall F(cF \equiv H \Rightarrow F)$, i.e., that c is a concept of *being human* just in case c encodes exactly the properties necessarily implied by *being human*. Thus, we would be able to derive that $ConceptOf(c,H) \rightarrow cE!$.

(**671**) **Theorems:** Existence Conditions for Concepts of Properties. In the usual way, it follows that (.1) there is a concept of G, (.2) there is a unique concept of G, and (.3) *the* concept of G exists:

(.1) $\exists c ConceptOf(c,G)$

(.2) $\exists! c ConceptOf(c,G)$

(.3) $\iota c ConceptOf(c,G)\downarrow$

(**672**) **Definition:** Notation for The Concept of G. We may now introduce the notation $\boldsymbol{c}_G$ for *the* concept of G:

$$\boldsymbol{c}_G =_{df} \iota c ConceptOf(c,G)$$

Thus, $\boldsymbol{c}_{()}$ is a term-forming operator that takes unary relation terms as arguments. The boldface symbol '$\boldsymbol{c}$' in the expression '$\boldsymbol{c}_G$' is *not* a variable ranging over concepts and, hence, not a restricted variable (if it were a variable, there would be a problem: it would occur free in the definiendum but not free in the definiens). The only variable in the expression '$\boldsymbol{c}_G$' is the symbol 'G', though

[339] By our conventions for restricted variables, the definition expands to:

$$ConceptOf(x,G) \equiv_{df} C!x \,\&\, G\downarrow \,\&\, \forall F(xF \equiv G \Rightarrow F)$$

Thus, for the definiendum to hold, both of the arguments of *ConceptOf* must be significant. So, by our conventions for definitions by equivalence, $ConceptOf(\kappa,\Pi)$ holds if and only if κ and Π both exist, κ is a concept, and κ encodes all and only the properties necessarily implied by Π.

we may regard '$\boldsymbol{c}_G$' itself as a functional term that takes a concept as a value depending on the value assigned to 'G'. And since the existence of G is built into the definition of $ConceptOf(c, G)$, then if Π is any empty unary relation term, $\boldsymbol{c}_\Pi$ is empty as well.

(**673**) **Theorem:** Identity of The Concept of G and The (Thick) Form of G. It should come as no surprise to those who have worked through the chapter on Forms that the concept of G is identical to the (thickly-conceived) Form of G:

$$\boldsymbol{c}_G = \Phi_G$$

This establishes an interesting link between the work of Plato and Leibniz. As an exercise, consider what the resulting theorems say when we substitute $\boldsymbol{c}_G$ for Φ_G in the theorems about Forms, and substitute Φ_G for $\boldsymbol{c}_G$ in the theorems below about concepts of properties.

(**674**) **Theorem:** The Concept of G is (Strictly) Canonical. By now familiar reasoning, $\boldsymbol{c}_G$ is (identical to) a canonical concept:

$$\boldsymbol{c}_G = \imath c \forall F(cF \equiv G \Rightarrow F)$$

Moreover, with theorem (448), we established, as a modally strict theorem, that for any property F, if G necessarily implies F, then it is necessary that G necessarily implies F, i.e., that $\forall F(G \Rightarrow F \rightarrow \Box G \Rightarrow F)$. Thus, where φ is $G \Rightarrow F$, this shows that φ is a rigid condition on properties, by (260.1). So $\boldsymbol{c}_G$ is (identical to) a strictly canonical concept, by (260.2).

(**675**) **Lemmas:** Facts About the Concept of G. Since $\boldsymbol{c}_G$ is identical to a strictly canonical concept, we may use (261.2) to establish, by modally strict means, (.1) for any property F, $\boldsymbol{c}_G$ encodes F if and only if G necessarily implies F:

(.1) $\forall F(\boldsymbol{c}_G F \equiv G \Rightarrow F)$

Moreover, since we can establish $G \Rightarrow G$ by modal predicate logic alone, it follows from (.1) that:

(.2) $\boldsymbol{c}_G G$

(**676**) **Theorem:** Identity of Sums. If we remember the definition of $\oplus$, then it is a simple consequence of the foregoing definitions that the sum of the concept of G and the concept of H is identical to the concept that encodes just the properties necessarily implied by G or necessarily implied by H:

$$\boldsymbol{c}_G \oplus \boldsymbol{c}_H = \imath c \forall F(cF \equiv G \Rightarrow F \vee H \Rightarrow F)$$

(**677**) **Metatheorems:** Reordering and Inclusion Chains. (.1) It is a theorem that no matter how one reorders the sum of the concepts $c_{G_1}, \ldots, c_{G_n}$, the original sum and the reordered sum are identical:

(.1) $\vdash \boldsymbol{c}_{G_1} \oplus \ldots \oplus \boldsymbol{c}_{G_{i-1}} \oplus \boldsymbol{c}_{G_i} \oplus \boldsymbol{c}_{G_{i+1}} \oplus \ldots \oplus \boldsymbol{c}_{G_{j-1}} \oplus \boldsymbol{c}_{G_j} \oplus \boldsymbol{c}_{G_{j+1}} \oplus \ldots \oplus \boldsymbol{c}_{G_n} =$
$\boldsymbol{c}_{G_1} \oplus \ldots \oplus \boldsymbol{c}_{G_{i-1}} \oplus \boldsymbol{c}_{G_j} \oplus \boldsymbol{c}_{G_{i+1}} \oplus \ldots \oplus \boldsymbol{c}_{G_{j-1}} \oplus \boldsymbol{c}_{G_i} \oplus \boldsymbol{c}_{G_{j+1}} \oplus \ldots \oplus \boldsymbol{c}_{G_n}$

for any $1 \leq i \leq j \leq n$. Moreover, it is a theorem that for any properties $G_1, \ldots, G_n$, the concept of G_1 is included in the sum of the concepts of G_1 and G_2, which in turn is included in the sum of the concepts of G_1, G_2, and G_3, ..., which in turn is included in the sum of the concepts G_1, G_2, ..., G_n; and the reverse is true for concept containment:

(.2) $\vdash \boldsymbol{c}_{G_1} \preceq \boldsymbol{c}_{G_1} \oplus \boldsymbol{c}_{G_2} \preceq \ldots \preceq \boldsymbol{c}_{G_1} \oplus \ldots \oplus \boldsymbol{c}_{G_n}$
$\vdash \boldsymbol{c}_{G_1} \oplus \ldots \oplus \boldsymbol{c}_{G_n} \succeq \ldots \succeq \boldsymbol{c}_{G_1} \oplus \boldsymbol{c}_{G_2} \succeq \boldsymbol{c}_{G_1}$

for any $n > 2$.

(**678**) **Remark:** Concepts and Properties. We have thus far analyzed Leibnizian concepts as abstract objects. But in Remark (622), we critically considered how the Leibnizian principles governing concept summation fare when concepts are instead analyzed as properties. We considered the definition:

$$F + G =_{df} [\lambda x\, Fx \,\&\, Gx]$$

and found that if F and G are properties (i.e., entities that can be distinct even though necessarily equivalent), then there is no obvious way to prove that summation is idempotent, commutative, and associative; see the discussion in Remark (622) for the details. In this Remark, we consider how the analysis of Leibnizian concepts as properties would give rise to additional difficulties when concept containment is brought into the picture and analyzed property-theoretically. And we examine a potential objection to the view that Leibnizian concepts are abstract objects.

Before we start our investigation, let's examine how the present analysis might be used to interpret Leibnizian texts. Earlier, in (629), we proved Leibniz's Equivalence in a completely general form, without reference to concepts of properties, for both concept inclusion and concept containment:

$$c \preceq d \equiv c \oplus d = d$$

$$c \succeq d \equiv c = c \oplus d$$

But if we instantiate these to concepts of properties, we obtain the following corollaries, one for concept inclusion and one for concept containment:

$$\boldsymbol{c}_F \preceq \boldsymbol{c}_G \;\equiv\; \boldsymbol{c}_F \oplus \boldsymbol{c}_G = \boldsymbol{c}_G$$

(ϑ) $\boldsymbol{c}_F \succeq \boldsymbol{c}_G \;\equiv\; \boldsymbol{c}_F = \boldsymbol{c}_F \oplus \boldsymbol{c}_G$

These are of interest because they more closely resemble Leibniz's texts. In Leibniz 1690b (LLP 131–132, G.vii 239), we find the following (in which the variables L and B have been replaced by G and F, respectively):

Proposition 13:
If $G + F = G$, then F will be in G

Proposition 14:
If F is in G, then $G + F = G$

Clearly, if we conjoin these and substitute $F + G$ for $G + F$ (which we can do since concept addition is commutative), then we obtain:

F is in G if and only if $F + G = G$

And the containment version is:

(ζ) F contains G if and only if $F = F + G$

On the present analysis of Leibnizian concepts, we've captured (ζ) as (ϑ).

However, if one wanted to suggest that instead of analyzing Leibnizian concepts as abstract individuals they should be analyzed as properties, then (ζ) raises another issue. For suppose that one offers the following property-theoretic analysis of concept containment:

$F\ contains\ G \ \equiv_{df}\ F \Rightarrow G$

Presumably, this definition would allow one to conclude that since *being a brother* necessarily implies *being male*, the property *being a brother* contains the property *being male*. So the definition would provide one with a way of understanding the Leibnizian claim that the concept *brother* contains the concept *male*.

But consider what happens when the above definition is coupled with the property-theoretic definition of concept summation discussed in Remark (622). Then (ζ) would be analyzed as:

(ξ) $F \Rightarrow G \equiv F = [\lambda x\, Fx \,\&\, Gx]$

(ξ), however, is inconsistent with reasonable principles about properties that one might wish to adopt. From the fact that property F necessarily implies property G, it does *not* follow that F just is identical to the conjunctive property $[\lambda x\, Fx \,\&\, Gx]$. One may consistently extend our theory with the claim:

$\exists F \exists G (F \Rightarrow G \ \&\ F \neq [\lambda x\, Fx \,\&\, Gx])$

For example, one may reasonably claim both (a) that the property *being a brother* necessarily implies the property *being male* yet (b) that *being a brother* is not identical to *being a brother and male*. Or if we let K be the necessary property $[\lambda x\, Fx \vee \neg Fx]$ (where F is any property), then one might reasonably argue that, for any property G distinct from K, both $G \Rightarrow K$ (indeed, this is provable) and $G \neq [\lambda y\, Gy \,\&\, Ky]$. Indeed, some philosophers might wish to go further and assert that the right condition of (ξ) fails universally, i.e., that no (distinct) properties F and G are such that F is identical to $[\lambda y\, Fy \,\&\, Gy]$. In any case, we only need the existence of a single counterexample to show that (ξ) is unacceptable.

Of course, one could try to reinterpret the identity claim in (ζ) in terms of some weaker notion. Castañeda (1990, 17) suggests that Leibniz's relation of coincidence is indeed a weaker relation than identity on concepts, but Ishiguro (1990, Chapter 2) argues that this isn't really consistent with Leibniz's texts. Among other things, it conflicts with Leibniz's reading of the symbol '∞' as 'the same' in Definition 1 of LLP 131 (G.vii 236). Of course, we've already ruled out Lenzen's (1990) interpretation of the variables in terms of sets, since this offers only a mathematical model of Leibnizian concepts, not a theory of them.

By contrast, we've seen that when the concept F is analyzed as $\boldsymbol{c}_F$, concept addition as $\oplus$, and concept containment as $\succeq$, then (ζ) becomes analyzed as (ϑ). The latter has the following instance:

$$\boldsymbol{c}_B \succeq \boldsymbol{c}_M \;\equiv\; \boldsymbol{c}_B = \boldsymbol{c}_B \oplus \boldsymbol{c}_M$$

The concept of *being a brother* contains the concept of *being male* if and only if the concept of *being a brother* is identical with the sum of the concepts of *being a brother* and *being male*.

Indeed, if one adds to our system the property-theoretic postulate that *being a brother* necessarily implies *being male* ($B \Rightarrow M$), one may derive either side of above biconditional.[340]

We conclude the present remark on concepts and properties by considering whether the following constitutes counterevidence to our analysis of concepts as abstract individuals. Some philosophers, including Leibniz, might assert that the following is true:

(A) The concept *brother* is identical to the sum of the concept *male* and the concept *sibling*.

Now in certain pre-theoretical, ordinary language contexts, in which *concept* talk is just loose talk about *properties* and concept summation is intuitively un-

[340] To see this, assume $B \Rightarrow M$, i.e., $\Box\forall x(Bx \rightarrow Mx)$. Now to derive the left-side of the biconditional, i.e., $\boldsymbol{c}_B \succeq \boldsymbol{c}_M$, we have to show $\boldsymbol{c}_M \preceq \boldsymbol{c}_B$, by definition (624.2). So by definition (624.1), we have to show $\forall F(\boldsymbol{c}_M F \rightarrow \boldsymbol{c}_B F)$. By GEN, it suffices to show $\boldsymbol{c}_M F \rightarrow \boldsymbol{c}_B F$. So assume $\boldsymbol{c}_M F$. Then by (675.1), $M \Rightarrow F$, i.e., $\Box\forall x(Mx \rightarrow Fx)$. But it follows from our assumption and this last fact that $\Box\forall x(Bx \rightarrow Fx)$, i.e., $B \Rightarrow F$. So by (675.1), $\boldsymbol{c}_B F$, completing our conditional proof. Now to derive the right-side of the biconditional from $B \Rightarrow M$, we could just appeal to the derivation we just completed and the instance of (ϑ) above in the text. Independently, however, we can easily show $\boldsymbol{c}_B = \boldsymbol{c}_B \oplus \boldsymbol{c}_M$ by showing they encode the same properties. ($\rightarrow$) Assume $\boldsymbol{c}_B F$. Then $\boldsymbol{c}_B F \vee \boldsymbol{c}_M F$. Hence $\boldsymbol{c}_B \oplus \boldsymbol{c}_M F$. ($\leftarrow$) Assume $\boldsymbol{c}_B \oplus \boldsymbol{c}_M F$. Then by (676), $(B \Rightarrow F) \vee (M \Rightarrow F)$. If $B \Rightarrow F$, then $\boldsymbol{c}_B F$, by (675.1). If $M \Rightarrow F$, then given our assumption $B \Rightarrow M$, it follows that $B \Rightarrow F$. Hence, again, $\boldsymbol{c}_B F$.

derstood as property conjunction, (A) is no doubt best interpreted as asserting that the property *being a brother* ('B') is identical to the property $[\lambda x\, Mx \,\&\, Sx]$, i.e., the conjunction of *being male* ('M') and *being a sibling* ('S'). This would yield a true interpretation of (A). But we have suggested in this subsection that Leibniz's talk of 'the concept brother' should be analyzed as a reference to *the concept of being a brother*, i.e., $\boldsymbol{c}_B$, and similarly for his talk of 'the concept male' and 'the concept sibling'. Thus one might be tempted to analyze (A) as:

(B) $\boldsymbol{c}_B = \boldsymbol{c}_M \oplus \boldsymbol{c}_S$

But (B) is provably false if given, as minimal facts about these properties, that *being a brother* necessarily implies *being a male sibling* (i.e., $B \Rightarrow [\lambda x\, Mx \,\&\, Sx]$), and that neither M nor S necessarily imply this conjunctive property. For, given such facts, it follows by definition of $\boldsymbol{c}_B$ and $\boldsymbol{c}_M \oplus \boldsymbol{c}_S$ that $\boldsymbol{c}_B$ encodes $[\lambda x Mx \& Sx]$ and that $\boldsymbol{c}_M \oplus \boldsymbol{c}_S$ does not. Since $\boldsymbol{c}_B$ encodes a property that $\boldsymbol{c}_M \oplus \boldsymbol{c}_S$ fails to encode, they are distinct, contradicting (B).

It may be tempting to conclude that this result shows that the property-theoretic analysis of the Leibnizian notion of concepts is preferable to the present analysis, since it avoids attributing to him a provable falsehood. As noted above, if we take the concepts discussed in (A) to be properties, and take the sum of the concepts *male* and *sibling* to be $[\lambda x\, Mx \,\&\, Sx]$, then (C) becomes an analysis of (A):

(C) $B = [\lambda x\, Mx \,\&\, Sx]$

(C) is in fact true. The property *being a brother* is identical to the property *being a male sibling*. Moreover, on the basis of (C), it follows that the concept of *being a brother* is identical to the concept of the conjunctive property *being a male sibling*:

(D) $\boldsymbol{c}_B = \boldsymbol{c}_{[\lambda x\, Mx \,\&\, Sx]}$

Once (C) is added to our system (say, as part of a theory of brotherhood), then (D) becomes an easy theorem.

So I think the conclusion to be drawn here is that either (A) is to be interpreted as (C), in which case Leibniz has asserted a property identity in the guise of concept identities, or (A) is to be interpreted as (D). If one doesn't rigorously distinguish properties and concepts of properties, one might well say (A) when intending (C) or (D). Leibniz did in fact distinguish between properties and their concepts, but it doesn't look like he did so rigorously. There are passages where he distinguishes an accident and its notion, and passages where he distinguishes a predicate and its concept. Consider the following passage from Article 8 in the *Discourse on Metaphysics* (PW 19, G.iv 433):

> ...an accident is a being whose notion does not include all that can be attributed to the subject to which this notion is attributed. Take, for example, the quality of being a king, ...

In this passage, I take it, Leibniz is distinguishing the property (i.e., the accident or the quality) from the concept or notion of that property. And recall the following passage from his correspondence with Arnauld (June 1686, LA 63, G.ii 56):

> ...in every true affirmative proposition, necessary or contingent, universal or particular, the concept of the predicate is in a sense included in that of the subject; the predicate is present in the subject.

These passages show Leibniz did distinguish properties and their concepts. But he didn't regiment the distinction and systematically adhere to it in his logic of concepts and containment theory of truth. So it shouldn't be surprising if sometimes Leibniz asserted concept identities like (A) when he meant to assert either something like (C) or (D). This conclusion may address the concern that (A) constitutes counterevidence to the present analysis.

Notice also that from (C), we can also derive the property-theoretic postulate discussed earlier, namely, (a) $B \Rightarrow M$, and so derive that (b) the concept of *being a brother* contains the concept of *being male*, i.e., that $\boldsymbol{c}_B \succeq \boldsymbol{c}_M$. For (a), note that $[\lambda x\, Mx \,\&\, Sx] \Rightarrow M$ (exercise) and so by (C), it follows that $B \Rightarrow M$. Hence, for (b), the reasoning we used in footnote 340 now yields the conclusion that $\boldsymbol{c}_B \succeq \boldsymbol{c}_M$.

In conclusion, on the present theory, there is a subtle and important difference between properties and concepts, between the sum concept $\boldsymbol{c}_M \oplus \boldsymbol{c}_S$ and the conjunctive property $[\lambda x\, Mx \,\&\, Sx]$, and hence between $\boldsymbol{c}_M \oplus \boldsymbol{c}_S$ and $\boldsymbol{c}_{[\lambda x\, Mx \,\&\, Sx]}$. In general, it is important to remember that adding concepts $\boldsymbol{c}_F$ and $\boldsymbol{c}_G$ to obtain $\boldsymbol{c}_F \oplus \boldsymbol{c}_G$ is not the same as conjoining properties F and G to obtain $[\lambda x\, Fx \,\&\, Gx]$. The conjunction $[\lambda x\, Fx \,\&\, Gx]$ is a property whose extension (in the technical sense of Chapter 10, Section 10.3) is the intersection of the extensions of F and G. The sum of $\boldsymbol{c}_F$ and $\boldsymbol{c}_G$ is a concept that encodes the union of the properties implied by F and those implied by G.

13.2.2 Concepts of Ordinary Individuals

(**679**) **Remark:** The Focus on Concepts of Ordinary Individuals. The notion of a *concept of an individual* immediately suggests the following general definition:

> x is a *concept of* y just in x is a concept and encodes exactly the properties that y exemplifies, i.e., $\mathit{ConceptOf}(x,y) \equiv_{df} C!x \,\&\, \forall F(xF \equiv Fy))$.

And given that it would be provable that, for every individual y, there is a unique such concept, we could then define:

$$\boldsymbol{c}_y \text{ (`}\textit{the}\text{ concept of } y\text{')} =_{df} \imath c ConceptOf(c, y)$$

However, there is a theoretical reason for not pursuing these general definitions. This concerns the fact that the properties that an abstract object exemplifies are not guaranteed to distinguish it from other abstract objects. Theorem (269) tells us that there are abstract objects x and y such that $x \neq y$ and $\forall F(Fx \equiv Fy)$. That is, there are distinct abstract objects (i.e., they encode different properties) that are indiscernible (i.e., exemplify the same properties). Given (269), it would then be straightforward to prove that there are distinct objects y and z whose concepts $\boldsymbol{c}_x$ and $\boldsymbol{c}_y$ are identical:

$$\exists x \exists y (x \neq y \,\&\, \boldsymbol{c}_x = \boldsymbol{c}_y)$$

So the concepts of distinct but indiscernible abstract individuals would provably collapse.

Of course, this reasoning doesn't apply either to concepts of *ordinary* individuals or to concepts of *discernible* individuals. Ordinary individuals are distinguished and identifiable in terms of the properties they exemplify (242.2), and so are discernible individuals (273.7). Since the *concepts* of such individuals encode the latter's exemplified properties, it makes sense to introduce concepts of ordinary individuals and, more generally, concepts of discernible individuals. Such a focus would ensure that if a and b are distinct ordinary individuals or distinct discernible individuals, their concepts $\boldsymbol{c}_a$ and $\boldsymbol{c}_b$ would be distinct.

However, there is a theoretical reason and various practical reasons for restricting our definitions, theorems, and proofs to concepts of ordinary objects. The theoretical reason stems from the fact that when we think about what concepts and abstract objects *are*, it seems reasonable to conclude that the very concept of an abstract object is given by the properties it *encodes*, not by the properties it exemplifies. So even if an abstract object x is discernible, the concept of x is, in its most fundamental sense, nothing other than x itself, since an abstract object encodes the properties by which it is *conceived* or *defined*. This constitutes a theoretical reason for not defining concepts of discernible objects: if the concept of a discernible abstract object x is intuitively just x, why introduce a notion of *concept* that distinguishes two objects when intuitively there is just one.

One practical reason for not defining concepts of discernible objects stems from the fact that our theory doesn't yet imply the existence of any discernible abstract objects. Later, in the chapter on natural numbers, we'll extend the theory with a new axiom, with the consequence that the extended theory does imply the existence of discernible abstract objects. But at present, only ordinary individuals are provably discernible; even if we allowed for concepts of discernibles, we wouldn't, at present, be able to prove that there are any concepts of discernible abstracta. So by restricting our attention to concepts of ordinary objects, we don't lose anything. Of course, one could argue that concepts of such discernibles might be of interest once we are in a position to prove they exist, the fact is that one can simply *reconstruct* the theorems below as theorems about discernibles should that prove to be of interest.

Moreover, by focusing on concepts of ordinary individuals, we will have ready-to-hand examples that Leibniz and others have used in discussing the metaphysics of individual concepts. When developing his calculus of individual concepts in the containment theory of truth and in his modal metaphysics, Leibniz always used concepts like *the concept of Adam*, *the concept of Sextus*, and *the concept of Alexander*. So a narrower focus on concepts of ordinary individuals comes with ready-made applications.

However, those who don't find the foregoing persuasive and who prefer to develop the metaphysics of concepts in the most general manner possible can easily revise the definitions, theorems, and proofs so that they govern concepts of discernible individuals. The proofs can be adjusted with little additional work. So it doesn't matter much which of these we study – there are some very nice results in the modal metaphysics of concepts in either case. These results, expressed in terms of concepts of ordinary individuals, are discussed in Section 13.4 below.

Given these observations, we'll use u, v as variables ranging over ordinary individuals in this subsection and for the remainder of this chapter. Since $O!x$ is a rigid restriction condition in the sense of (340), we may reason with u, v as free restricted variables and employ our extended Rule RN (341).

(**680**) **Definition:** Concepts of Ordinary Individuals. If we employ our theory of definitions by equivalence and conventions for free restricted variables, we may say that c is a *concept of* u just in case c encodes exactly the properties that u exemplifies:

$$\mathit{ConceptOf}(c,u) \equiv_{df} \forall F(cF \equiv Fu)$$

To take a simple example, if Socrates ('s') is an ordinary object and a concept c encodes exactly the properties Socrates exemplifies, then $\mathit{ConceptOf}(c,s)$.[341]

[341] Note again that by using the free restricted variables c and u, our rules and Convention (338.2) imply that the definition yields the following as a theorem:

$$\mathit{ConceptOf}(x,y) \equiv (C!x \,\&\, O!y \,\&\, \forall F(xF \equiv Fy))$$

Moreover, c and u are *rigid* restricted variables, since $\vdash_{\Box} \forall x(C!x \to \Box C!x)$ and $\vdash_{\Box} \forall y(O!y \to \Box O!y)$. So we may treat $C!c$ and $O!u$ as necessary axioms. Thus it follows from the preceding equivalence that:

(**681**) **Theorems:** Existence Conditions for Concepts of Ordinary Individuals. It follows that (.1) there is a a concept of u; (.2) there is a unique concept of u; and (.3) the concept of u exists:

(.1) $\exists c ConceptOf(c,u)$

(.2) $\exists! c ConceptOf(c,u)$

(.3) $\iota c ConceptOf(c,u)\downarrow$

These are straightforward consequences of the definition (680) and comprehension for concepts (614).

Of course, in these theorems, u is a free restricted variable and so, strictly speaking, all of the above are conditionals. But we can derive *unconditional* existence claims from them since it follows from (227.1) and the T schema that ordinary individuals exist.

(**682**) **Definition:** Notation for the Concept of u. We henceforth use the notation $\boldsymbol{c}_u$ to denote the concept of u:

$$\boldsymbol{c}_u =_{df} \iota c ConceptOf(c,u)$$

Again, the boldface expression '$\boldsymbol{c}$' in $\boldsymbol{c}_u$ is not a free variable and so not a restricted free variable; rather it is part of the functional expression $\boldsymbol{c}_u$, whose only free variable is u. So if we extend our language with the constant 's', to designate a known ordinary object, say Socrates, then $\boldsymbol{c}_s$ is the concept of Socrates.

It is important to recognize that both methods of eliminating the free restricted variable u and bound restricted variable c in the above, by using our conventions in (337), (338), and (339), yield the same result. On the one hand, we could start with the above definition as given, eliminate the free restricted variable u as outlined in (339), so that the definition becomes:

$$\boldsymbol{c}_y =_{df} \iota c(O!y \,\&\, ConceptOf(c,y))$$

and then eliminate the bound restricted variable in the definiens, as described in (337), so that c_y becomes defined as $\iota x(C!x \,\&\, O!y\; ConceptOf(x,y))$. Alternatively, we could start by eliminating the two free restricted variables in the definition of $ConceptOf(c,u)$ (680), as described in (338), so that the definition becomes:

$$ConceptOf(x,y) \equiv_{df} C!x \,\&\, O!y \,\&\, \forall F(xF \equiv Fy)$$

$$ConceptOf(c,u) \equiv \forall F(cF \equiv Fu)$$

The derivability of this equivalence justifies the definition given in the text.

Then we could introduce $\boldsymbol{c}_y$ via the definiens $\iota x(C!x \,\&\, O!y\, \mathit{ConceptOf}(x,y))$.

In either case, theorem (681.2) ensures that for every ordinary object y, $\exists! c\mathit{ConceptOf}(c,y)$ and theorem (681.3) ensures that $\iota c(\mathit{ConceptOf}(c,y){\downarrow}$. So $\boldsymbol{c}_u{\downarrow}$, for every u.

(**683**) **Theorem:** Contingency of Exemplification for Ordinary Objects. With the help of a previous theorem (205.1), we now prove that there exists an ordinary object u and a property F such that it is possible that (a) u exemplifies F and (b) possibly u fails to exemplify F:

$$\exists u \exists F \Diamond(Fu \,\&\, \Diamond\neg Fu)$$

By (165.11), this implies that for some ordinary object u and property F, $\Diamond Fu \,\&$ $\Diamond\neg Fu$. So there are objects u and properties F such that x possibly exemplifies F and possibly does not.

Exercise. Is it provable that *every* ordinary object u is such that for some property F, both $\Diamond Fu$ and $\Diamond\neg Fu$, i.e., $\forall u \exists F(\Diamond Fu \,\&\, \Diamond\neg Fu)$? Consider the fact that the claim $\exists x \Box E!x$, i.e., there exists an object that is necessarily concrete (e.g., an object such as Spinoza's God), seems to be consistent with the theory. But also consider the fact that for any two possible worlds w and w', there is at least one proposition, say p, and thus one property, $[\lambda y\, p]$, that distinguishes them.

(**684**) **Remark:** The Concept of u is Not Strictly Canonical. As noted earlier, we know that both $\iota c(\mathit{ConceptOf}(c,u){\downarrow}$ and $\boldsymbol{c}_u{\downarrow}$, for every u. So whenever u is a ordinary object, it follows, by now familiar reasoning, that:

$$\boldsymbol{c}_u = \iota x(A!x \,\&\, \forall F(xF \equiv Fu))$$

So $\boldsymbol{c}_u$ is canonical object.

But it is not difficult to see that $\boldsymbol{c}_u$ is not a strictly canonical object. Recall that in (260.2), we stipulated that $\iota x(A!x \,\&\, \forall F(xF \equiv \varphi))$ is a *strictly* canonical object just in case φ is a rigid condition on properties, i.e., by (260.1), just in case $\vdash_{\Box} \forall F(\varphi \to \Box\varphi)$. If we let φ be the formula Fu, then theorem (683) has the form $\exists u \exists F \Diamond(\varphi \,\&\, \Diamond\neg\varphi)$. Then by reasoning analogous to that in Remarks (298) and (326), it follows that φ is not a rigid condition on properties, on pain of system inconsistency. So though $\boldsymbol{c}_u$ is (identical to) a canonical object, it is not (identical to) a strictly canonical object.

This is, of course, as it should be. Our system doesn't assert, of any particular ordinary individual u and property F, that $\Diamond Fu \,\&\, \Diamond\neg Fu$. But it does imply the existence of ordinary individuals. To establish that $\boldsymbol{c}_u F$ when F is a contingent property, we shall need to appeal to the contingent fact that Fu, and this undermines the modal strictness of the reasoning. Thus, we should not expect completely general theorems about the properties $\boldsymbol{c}_u$ encodes to be modally strict.

(**685**) ⋆**Lemmas:** Fact About Concepts of Ordinary Objects. An immediate consequence of the foregoing is that the concept of u encodes G iff u exemplifies G:

$$\boldsymbol{c}_u G \equiv Gu$$

This theorem is not modally strict. **Exercise**: Show that $\boldsymbol{c}_u G \equiv \mathcal{A}Gu$ is a modally strict theorem.

(**686**) **Lemma:** Fact Relating Two Kinds of Concepts. It follows by modally strict means that the concept of u encodes G iff the concept of u contains the concept of G:[342]

$$\boldsymbol{c}_u G \equiv \boldsymbol{c}_u \succeq \boldsymbol{c}_G$$

(**687**) **Definition:** Completeness of Concepts. Recall that $\overline{F}$ was defined in (196.1) to be $[\lambda y\, \neg Fy]$. We now say that a concept c is *complete* just in case for every property F, either c encodes F or c encodes $\overline{F}$:

$$Complete(c) \;\equiv_{df}\; \forall F(cF \vee c\overline{F})$$

(**688**) **Theorem:** The Concept of (Ordinary) Individual u Is Complete.

$$Complete(\boldsymbol{c}_u)$$

This theorem and (685)⋆ capture Leibniz's suggestion, in Article 8 of the *Discourse on Metaphysics* (PW 18–19, G.iv 433), that:

> ...it is in the nature of an individual substance, or complete being, to have a notion so complete that it is sufficient to contain and render deducible from itself, all the predicates of the subject to which this notion is attributed.

Theorem (685)⋆ ensures that all of the properties that u exemplifies are properties that $\boldsymbol{c}_u$ encodes, and the present theorem ensures that $\boldsymbol{c}_u$ is provably complete. But to capture Leibniz's containment theory of truth, we'll have to derive a further fact from (685)⋆ and (686). See theorem (692.1)⋆ and Remark (693) below, where we show that the true affirmative judgment Fu implies that the *concept* of G is contained in the *concept* of u, and vice versa.

13.2.3 Concepts of Generalizations

In what follows, we introduce concepts of *generalizations*, such as the concept *everything that exemplifies G* and the concept *something that exemplifies G*. In our metalanguage, we take the liberty of rendering these expressions more simply

[342] Thanks to Daniel West for noting that the following is provable by modally strict means.

as the concept *every* G and the concept *some* G, and we introduce and define notation like $\boldsymbol{c}_{\forall G}$ and $\boldsymbol{c}_{\exists G}$, respectively, to represent them. To further simplify the discussion, these formal expressions will be defined directly in terms of canonical descriptions.

(**689**) **Definitions:** Concepts of Generalizations. Let us say (.1) the concept *everything that exemplifies* G (i.e., the concept *every* G) is the concept that encodes just the properties F such that every G exemplifies F; and (.2) the concept *something that exemplifies* G (i.e., the concept *some* G) is the concept that encodes just the properties F such that some G exemplifies F:

(.1) $\boldsymbol{c}_{\forall G} =_{df} \iota c\forall F(cF \equiv \forall x(Gx \to Fx))$

(.2) $\boldsymbol{c}_{\exists G} =_{df} \iota c\forall F(cF \equiv \exists y(Gy \,\&\, Fy))$

In these definitions, the variables x, y are unrestricted and range over individuals generally. Thus, $\boldsymbol{c}_{\forall G}$ and $\boldsymbol{c}_{\exists G}$ are canonical objects. We leave it as an exercise to show that $\boldsymbol{c}_{\forall G}$ and $\boldsymbol{c}_{\exists G}$ are not strictly canonical, i.e., that when φ is either $\forall x(Gx \to Fx)$ or $\exists x(Gx \,\&\, Fx)$, then φ is *not* a rigid condition on properties.

(**690**) ⋆**Lemmas:** Facts About Concepts of Generalizations. It is an immediate consequence of the foregoing that (.1) the concept *every* G encodes F iff every G exemplifies F, and (.2) the concept *some* G encodes F iff some G exemplifies F:

(.1) $\boldsymbol{c}_{\forall G}F \equiv \forall x(Gx \to Fx)$

(.2) $\boldsymbol{c}_{\exists G}F \equiv \exists x(Gx \,\&\, Fx)$

(**691**) **Lemmas:** Modally Strict Facts About Concepts of Generalizations. By modally strict means, we may prove that (.1) the concept *every* G encodes the property F iff the concept *every* G contains the concept of F, and (.2) the concept *some* G encodes the property F iff the concept *some* G contains the concept of F:

(.1) $\boldsymbol{c}_{\forall G}F \equiv \boldsymbol{c}_{\forall G} \succeq \boldsymbol{c}_F$

(.2) $\boldsymbol{c}_{\exists G}F \equiv \boldsymbol{c}_{\exists G} \succeq \boldsymbol{c}_F$

13.3 The Containment Theory of Truth

(**692**) ⋆**Theorems:** Exemplification and Containment. It is a consequence of the preceding definitions and lemmas that exemplification predication is equivalent to the Leibnizian analysis of predication, for it follows that (.1) u exemplifies G if and only if the concept of u contains the concept of G. Furthermore, the modern and Leibnizian analyses of simple quantified statements are equivalent, given the following theorems: (.2) everything exemplifying G exemplifies F if and only if the concept *everything that exemplifies* G contains the concept of F; and (.3) something exemplifies both G and F if and only if the concept *something that exemplifies* G contains the concept of F:

(.1) $Gu \equiv \boldsymbol{c}_u \succeq \boldsymbol{c}_G$

(.2) $\forall x(Gx \rightarrow Fx) \equiv \boldsymbol{c}_{\forall G} \succeq \boldsymbol{c}_F$

(.3) $\exists x(Gx \,\&\, Fx) \equiv \boldsymbol{c}_{\exists G} \succeq \boldsymbol{c}_F$

These are easy consequences of previous theorems.

(**693**) **Remark:** The Containment Theory of Truth. To see how Leibniz's containment theory of truth is preserved in these theorems, first recall the passage quoted at the outset, from the correspondence with Arnauld (June 1686, LA 63, G.ii 56):

> ... in every true affirmative proposition, necessary or contingent, universal or particular, the concept of the predicate is in a sense included in that of the subject; the predicate is present in the subject.

Leibniz also produced an early statement of his containment theory of truth in a work subsequently titled *Elements of a Calculus*, where he spoke of universal propositions and wrote (1679, 18–19; source C 51):

> ... every true universal affirmative categorical proposition simply shows some connection between predicate and subject (a *direct* connection, which is what is always meant here). This connection is, that the predicate is said to be in the subject, or to be contained in the subject; either absolutely and regarded in itself, or at any rate, in some instance; i.e., that the subject is said to contain the predicate in a stated fashion. This is to say that the concept of the subject, either in itself or with some addition, involves the concept of the predicate....

In Article 8 of the *Discourse on Metaphysics* (1686, Bennett's translation of G.iv 433) we find:[343]

> So the [notion of the] subject term must always include [that of] the predicate, so that anyone who understood the subject notion perfectly would also judge that the predicate belongs to it.

Now to see how these passages are validated by our theorems, let a stand for Alexander the Great and let us assume Alexander is an ordinary object. Then the concept of Alexander, $\boldsymbol{c}_a$, is significant, by (682) and (681.3). Moreover, let K stand for the property *being a king*. Then the concept of *being a king*, $\boldsymbol{c}_K$, is well-formed and logically proper, by (672) and (671.3). Now, given the above quotations, Leibniz analyzes the ordinary language predication:

[343] The words in brackets were interpolated by Bennett, so as to clarify Leibniz's intention.

(.1) Alexander is a king

in the following terms:

(.2) The concept of Alexander contains the concept of *being a king*

On our representation of Leibniz's analysis, this becomes:

(.3) $\boldsymbol{c}_a \succeq \boldsymbol{c}_K$

Whereas the modern analysis of (.1) in the predicate calculus is:

(.4) Ka

So by (692.1)★, our representation (.3) of Leibniz's analysis (.2) of the ordinary claim (.1) is equivalent to the modern analysis (.4) of (.1).

As a second example, consider the ordinary language predication:

(.5) Every person is rational.

Leibniz's analysis would be:

(.6) The concept *every person* contains the concept *being rational*

We represented the Leibnizian concept *every person* as the concept *everything that exemplifies being a person*, i.e., $\boldsymbol{c}_{\forall P}$, which is defined in (689.1). So our representation of Leibniz's analysis (.6) becomes:

(.7) $\boldsymbol{c}_{\forall P} \succeq \boldsymbol{c}_R$

Whereas the modern analysis of (.5) in the predicate calculus is:

(.8) $\forall x(Px \rightarrow Rx)$

So by (692.2)★, our representation (.7) of Leibniz's analysis (.6) of the ordinary claim (.5) is equivalent to the modern analysis (.8) of (.6). And analogously for the ordinary language predication 'Some person is rational'.

(**694**) **Remark:** Generalized Quantifiers. Is is intriguing to consider whether there is a connection between Leibniz's containment theory of truth and the idea of a generalized quantifier or Montague's (1974) subject-predicate analysis of basic sentences of natural language. Montague gave a uniform subject-predicate analysis of a fundamental class of English sentences by treating such noun phrases as 'John' and 'every person' as sets of properties. He supposed that the proper name 'John' denotes the set of properties that John exemplifies and supposed that the noun phrase 'every person' denoted the set of properties that every person exemplifies. Then, on Montague's theory, English sentences such as 'John is rational' and 'Every person is rational' could be given a uniform subject-predicate analysis: such sentences are true iff the property denoted by the predicate 'is rational' is a member of the set of properties denoted by the subject term. Leibniz's containment theory of truth, on our analysis, offers a similar unification of the analysis of singular and general predications. Moreover, one could extend this to other concepts of generalizations, such as the concepts *most things that exemplify F*, *many things that exemplify F*, *few things that exemplify F*, etc., once the relevant quantifiers are introduced into our language, either as axiomatized primitives or by definition.

(**695**) **Remark:** Hypothetical Necessity. Leibniz's containment theory of truth gives rise to an interesting objection. Leibniz anticipated the objection in Article 13 of the *Discourse on Metaphysics*. In that Article, Leibniz reiterates that "the notion of an individual substance involves, once and for all, everything that can ever happen to it", and then says (PW 23, G.iv 436):

> But it seems this will destroy the difference between contingent and necessary truths, that human freedom will no longer hold, and that an absolute fatality will rule over all our actions as well as over the rest of what happens in the world.

Leibniz's contemporary, Antoine Arnauld, took this criticism in a theistic direction, suggesting that the view not only implies that everything that happens to a person happens by necessity but also that it places constraints on God's freedom to shape what happens to the history of the human race (letter to Count Ernst von Hessen-Rheinfels, March 13, 1686, LA 9, G.ii 15).

Let us put aside Arnauld's theistic turn. Given Leibniz's phrasing, the objection charges that the containment theory of truth somehow collapses contingency and necessity and that, at best, the theory *represents* contingent truths as necessary truths and, at worst, implies that the actual world exhibits no contingency. If concept containment is not relative to any circumstance, then Leibniz has analyzed the contingent statement 'Alexander is a king' in terms of a claim that appears to be a necessary truth, namely, the concept of Alexander contains the concept of *being a king*. A similar worry arises about the analysis of the contingent general claims 'Every person is rational' and 'Some person is rational', which if true, would also appear to be rendered as necessary truths given Leibniz's analysis. For the purposes of simplying the discussion, we'll put these last two examples aside and focus just on the first, since the discussion of the first example will apply to them as well.

Leibniz's response to the objection is somewhat intriguing. In the next passage of the *Discourse*, he continues (PW 23-24, G.iv 437):

> To give a satisfactory answer to it, I assert that connexion or sequence is of two kinds. The one is absolutely necessary, whose contrary implies a contradiction; this kind of deduction holds in the case of eternal truths, such as those of geometry. The other is only necessary by hypothesis (*ex*

> *hypothesi*), and so to speak by accident; it is contingent in itself, since its contrary does not imply a contradiction.

So Leibniz defends himself against the objection by appealing to a distinction between absolute necessity and hypothetical necessity. The question is, what is hypothetical necessity?

Before we propose an answer to this question, note that Jonathan Bennett offers a slightly different translation of the passage from the *Discourse* just quoted. In his translation, Bennett (a) makes it much clearer that Leibniz is talking about two kinds of *logical consequence*, and (b) interpolates some text that helps us to understand what Leibniz is suggesting. Bennett's translation goes as follows, with the added material enclosed between center dots:

> To that end, I remark that there are two kinds of connection or following from. One is absolutely necessary, and its contrary implies a contradiction; such deduction pertains to eternal truths, such as those of geometry. The other is necessary ·not absolutely, but· only *ex hypothesi*, and, so to speak, accidentally. ·It doesn't bring us to *It is necessary that P*, but only to *Given Q, it follows necessarily that P*·. Something that is necessary only ex hypothesi is contingent in itself, and its contrary doesn't imply a contradiction.

Now I'm not sure why Bennett interpolated the text between the dots in his translation;[344] it doesn't appear in the original at G.iv 437. But his explanation of the distinction between absolute and hypothetical necessity seems to be on point, though it introduces a crucial ambiguity.

This translation leaves it open whether 'necessarily' in the phrase 'Given Q, it follows necessarily that P' attaches to 'it follows that' or to the proposition P. If the former, then Leibniz is thinking of hypothetical necessities in terms of a more formal, inferential notion of logical consequence, in which the conclusion follows by logical necessity from accidental (i.e., contingent) truths. But if the 'necessarily' in '... it follows necessarily that P' attaches to the proposition P, then we would have to find an interpretation on which a necessary truth becomes derivable from a contingency. Let's consider these interpretative options in turn.

Suppose that the 'necessarily' attaches to 'it follows that', so that the necessity in question is logical necessity, i.e., a conclusion derived from a contingent hypothesis is *logically required*. We can explore this idea by focusing either on derivations or on theorems. Focusing first on derivations, our system establishes that $Ka \vdash \boldsymbol{c}_a \succeq \boldsymbol{c}_K$. Theorem (692.1)★ asserts $Ka \equiv \boldsymbol{c}_a \succeq \boldsymbol{c}_K$, which implies

[344] In a note at the top of the translation, Bennett writes:

> Small ·dots· enclose material that has been added, but can be read as though it were part of the original text.

This doesn't say what the origin of the added material is.

$Ka \rightarrow \boldsymbol{c}_a \succeq \boldsymbol{c}_K$. Thus, by (63.10), $Ka \vdash \boldsymbol{c}_a \succeq \boldsymbol{c}_K$. So Leibniz's containment analysis of 'Alexander is a king', is derivable from the contingent premise Ka; it is a conclusion that is a logically-required consequence of a contingent premise. This is one way to understand hypothetical necessity.

But if we focus on theorems instead of derivations, then $Ka \equiv \boldsymbol{c}_a \succeq \boldsymbol{c}_K$ is a theorem (692.1)★ that is a logical consequence of our axioms and rules, i.e., those axioms and rules logically require it to be a theorem. We may not conclude that it is a necessary truth, since it has been established by non-modally strict means and ultimately rests on a modally fragile axiom. So this is a second way to understand hypothetical necessity.

Now, let's consider the other interpretative option, in which 'necessarily' in the phrase 'Given Q, it follows necessarily that P' attaches to the proposition P. Then we may regard such hypothetical necessities as metaphysically necessary truths derivable from contingent premises. An example is found in the claim that the concept of Alexander encodes *being a king*, i.e., $\Box\boldsymbol{c}_aK$. In object theory it is easy to show $Ka \vdash \Box\boldsymbol{c}_aK$. To see that this holds, note that $Ka \vdash \boldsymbol{c}_aK$, by (685)★ and (63.10). That is, the premise that Alexander is king implies that the concept of Alexander encodes *being a king*. But by axiom (51) and (63.10), we know can conclude $\boldsymbol{c}_aK \vdash \Box\boldsymbol{c}_aK$. Hence, from these two facts about derivations, it follows by (63.8) that $Ka \vdash \Box\boldsymbol{c}_aK$. So a necessary truth, $\Box\boldsymbol{c}_aK$, is derivable from a contingent premise. This is a third way to understand hypothetical necessity.

On none of these ways of understanding hypothetical necessity does it follow that the Leibnizian containment theory of truth banishes contingency from the actual world. But one might still wonder whether Leibniz and Arnauld might have made the objection stronger by recasting it as follows: the equivalence of an analysandum and its analysans should be a necessary equivalence, but the equivalence of the analysandum 'Alexander is a king' (Ka) and the analysans $\boldsymbol{c}_a \succeq \boldsymbol{c}_K$ is not necessary. The fact being appealed to in this objection is correct, for as we know, the equivalence $Ka \equiv \boldsymbol{c}_a \succeq \boldsymbol{c}_K$, is an instance of theorem (692.1)★ and so not a modally strict theorem. So we can't validly derive $\Box(Ka \equiv \boldsymbol{c}_a \succeq \boldsymbol{c}_K)$ by applying RN to (692.1)★.

But though it is understandable why one might raise this objection, it can be met. For it is not clear why an analysans should be necessarily equivalent to its analysandum in a system in which all terms rigidly designate and abstractions can be defined on the basis of contingencies. In the present system, $\boldsymbol{c}_a$ is defined on the basis of the properties Alexander in fact exemplifies, some of which are contingently exemplified. It is therefore inevitable that claims about the properties $\boldsymbol{c}_a$ (rigidly) encodes become provably, but not necessarily, equivalent to claims about the properties Alexander in fact exemplifies. Thus, it may be that a correct analysis requires only that the equivalence of the analysans and analysandum be provable *a priori* rather than provably necessary.

13.4 The Modal Metaphysics of Concepts

(**696**) **Remark:**. Primitive vs. Defined Counterparts. The theorems described in this section articulate a modal metaphysics of concepts inspired by Leibniz's work. We may introduce these theorems by way of an issue in Leibniz scholarship. Some Leibniz scholars have suggested that the best way to reconstruct Leibniz's modal metaphysics of concepts is to suppose that the *counterpart* relation partially systematized in Lewis 1968 should be applied to individual concepts. This general view is adopted here, but instead of taking *counterpart* to be a primitive, as in Lewis 1968 and in the work of other Leibniz scholars, we shall define a notion of *individual concept* and define the conditions under which one individual concept is a *counterpart of* another. Since our system axiomatizes a fixed domain of individuals and presupposes that every individual exemplifies properties in every possible world, it turns out that the modal metaphysics will preserve elements of Leibnizian, Lewisian, and Kripkean metaphysics.

The view adopted by various Leibniz scholars, that one should use counterpart theory to reconstruct Leibniz's metaphysics, traces back to work of Mondadori (1973, 1975), who notes that the natural reading of certain passages in the Leibnizian corpus are suggestive of that theory (cf. Ishiguro 1972, 123–134). Here is a passage from the *Theodicy* (T 371, G.vi 363) which Mondadori cites:

> I will now show you some [worlds], wherein shall be found, not absolutely the same Sextus as you have seen (that is not possible, he carries with him always that which he shall be) but several Sextuses resembling him, possessing all that you know already of the true Sextus, but not that is already in him imperceptibly, nor in consequence all that shall yet happen to him. You will find in one world a very happy and noble Sextus, in another a Sextus content with a mediocre state, . . .

Mondadori also cites the letter to Count Ernst von Hessen-Rheinfels of April 12, 1686, where Leibniz talks about the different possible Adams, all of which differ from each other (PW 51, G.ii 20):

> For by the individual notion of Adam I undoubtedly mean a perfect representation of a particular Adam, with given individual conditions and distinguished thereby from an infinity of other possible persons very much like him, but yet different from him. . . There is one possible Adam whose posterity is such and such, and an infinity of others whose posterity would be different; is it not the case that these possible Adams (if I may so speak

> of them) are different from one another, and that God has chosen only one of them, who is exactly our Adam?

When Mondadori suggests using counterpart theory to model Leibniz's views, he notes that whereas for Lewis the counterpart relation is a relation on individuals, "in Leibniz's case, it is best regarded as being a relation between (complete) concepts" (1973, 248). This suggestion is explicitly built into the Leibnizian system described in Fitch 1979, and has been adopted by other commentators as well.[345]

According to this view, in a Leibnizian modal metaphysics, a possible world is not a locus of compossible individuals but rather a locus of *compossible individual concepts*. A reconstruction of such metaphysics would involve definitions that, *intuitively*, (1) induce a partition on the domain of individual concepts into equivalence classes of compossible individual concepts, (2) induce a one-to-one correspondence between groups of compossible individual concepts and the possible worlds where they 'appear', and (3) induce a separate partition of the domain of individual concepts into equivalence classes of *counterpart* individual concepts, which groups an individual concept at one possible world with its counterparts at other possible worlds. Such a reconstruction has to ensure that an ordinary claim such as:

> Alexander is a king but might not have been.

becomes equivalent to the following claim:

> The concept of Alexander contains the concept of *being a king*, but there is an individual concept c such that: (i) c is a counterpart of the concept of Alexander, (ii) c doesn't contain the concept of *being a king*, and (iii) c appears at some other possible world.

In our reconstruction of these ideas below, this equivalence is preserved as a fundamental theorem of Leibniz's modal metaphysics, as item (736.1)⋆.

Now when Leibniz talks about the 'many possible Adams' and 'several Sextuses' that are all distinct from one another, the above-mentioned commentators take him to be talking about different concepts of the same individual rather than different possible individuals. In the modal metaphysics developed below, this suggestion is preserved, but we do *not* stipulate that the different concepts of Adam that appear at the various possible worlds stand in a primitive counterpart relation. Instead, we begin with the Kripkean view that Adam

[345] See Wilson 1979 and Vailati 1986. Lloyd (1978) also accepts that Leibniz 'resorts to counterparts' (p. 379), though she discovers some Leibnizian features in a Kripkean semantics of rigid designators, which assumes that the same individual can appear in other possible worlds. Interestingly, Kripke notes that "Many have pointed out to me that the father of counterpart theory is probably Leibnitz [sic]" (1980, 45, n. 13).

himself exemplifies different properties at different possible worlds. We then examine the different concepts of Adam that these exemplification facts *induce* on the domain of abstract objects. On the basis of this structure, we can define the sense in which the various concepts of Adam constitute counterparts of one another.

It is also important to remember again that our definitions will be cast within the context of the simplest quantified modal logic, in which there is a single, fixed domain of individuals. Thus, not only do the properties that Adam exemplifies at one possible world differ from the properties he exemplifies at other possible worlds, the ordinary individual Adam exemplifies properties at *every* possible world. For example, although Adam is concrete at our world and at certain other possible worlds, there are possible worlds where he fails to be concrete. In many of the possible worlds where Adam is concrete, his 'posterity is different'. At possible worlds where Adam is not concrete, he has no posterity.

Now although the same ordinary individual Adam exemplifies properties at every other possible world, a Leibnizian metaphysics of 'world-bound' abstract individuals emerges once we consider, for each possible world w, the concept that encodes exactly the properties that Adam exemplifies at w. At each possible world, Adam realizes a *different* concept, since concepts differ whenever they encode distinct properties. The concept that encodes all and only the properties Adam exemplifies at one possible world is distinct from the concept that encodes all and only the properties Adam exemplifies at a different possible world, though all of the different concepts of Adam will be counterparts of one another. Of course, we will define only one of these concepts to be *the concept of Adam*, namely, the concept that encodes just what Adam exemplifies at the actual world. In other words, once we relativize concepts of Adam to a world w, then *the* concept of Adam will be identified with the concept of Adam at $\boldsymbol{w}_\alpha$. When Leibniz talks about 'possible Adams', we may take him to be talking about different Adam-at-w concepts. We'll explain this in more detail once the definitions and theorems have been presented.

What is more interesting is the fact that these individual concepts have certain other Leibnizian features. We shall not just *stipulate* that compossibility is an equivalence condition on individual concepts, but rather *define* compossibility and *prove* that it is such a condition. Moreover, we shall not define possible worlds as sets of compossible individual concepts, but rather prove that there is a one-to-one correspondence between the groups of compossible concepts and the possible worlds. We shall also show that there is a sense of *mirrors* for which it is provable that each member of a group of compossible individual concepts mirrors its corresponding possible world.

So, although we do not use counterpart-theoretic primitives in our reconstruction, we nevertheless *recover* a modal metaphysics in which complete individual concepts are *world-bound* (in the sense of appearing at a unique world). These individual concepts will provide an interpretation for much of what Leibniz says about necessity, contingency, completeness, mirroring, etc. and, in the process, offer a way to reconcile Kripke's and Lewis's modal metaphysics.

13.4.1 Realization, Appearance, Mirroring

In this subsection, we continue to use the variables u, v to range over ordinary objects, and use w to range over possible worlds, as defined in Chapter 12.

(**697**) **Definition:** Realization at a World. Making use of the notion of *truth at* a possible world ($w \models p$), as defined in (515) and (470), let us say that an ordinary object u *realizes* a concept c *at* possible world w just in case for all properties F, the proposition u-exemplifies-F is true at w if and only if c encodes F:

$$RealizesAt(u,c,w) \equiv_{df} \forall F(w \models Fu \equiv cF)$$

In other words, u realizes c at w just in case u exemplifies at w exactly the properties c encodes.

(**698**) **Remark:** In what follows, we shall investigate notions definable, and theorems expressible, in terms of *RealizesAt*. This will be a completely general study; we won't prove any particular claims of the form $RealizesAt(u,c,w)$. Indeed, we can't do so, and the reason is that our system hasn't yet been applied and so doesn't identify any particular ordinary objects. We do know, however, that there are ordinary objects.

To be maximally explicit, note that the following is provable, though by non-modally strict means:

$$\exists u \exists c \exists w RealizesAt(u,c,w)$$

Proof. By (227.1) and the T-schema, $\exists x O!x$. Let a be such an object, so that we know $O!a$. Now consider $\boldsymbol{c}_a$, which clearly exists. We now show that a, $\boldsymbol{c}_a$, and $\boldsymbol{w}_\alpha$ are witnesses to our claim. So by (697), we have to show $\forall F(\boldsymbol{w}_\alpha \models Fa \equiv \boldsymbol{c}_a F)$ or, by GEN, $\boldsymbol{w}_\alpha \models Fa \equiv \boldsymbol{c}_a F$. To do this, note that by (685)★, we know $\boldsymbol{c}_a F \equiv Fa$. Moreover, by (536)★, we know $Fa \equiv \boldsymbol{w}_\alpha \models Fa$. Hence, $\boldsymbol{c}_a F \equiv \boldsymbol{w}_\alpha \models Fa$ and, by the symmetry of the biconditional, $\boldsymbol{w}_\alpha \models Fa \equiv \boldsymbol{c}_a F$.

But though we can prove this claim, we can't prove, for any particular ordinary individual, that it realizes a particular concept at a world. Our unapplied theory doesn't identify or individuate any particular ordinary individual. So, the reader should recognize that though the theorems in what follows do govern ordinary objects, concepts, and worlds, they are, in some sense, *projective*. Of course, on occasion, it will be helpful to illustrate these claims with intuitive examples. But such examples presuppose that the theory has been appropriately extended with new (and sometimes contingent) facts.

(**699**) **Theorem:** Facts About Realization. The definition of realization at a world has the following consequences. (.1) if, for some ordinary object u and possible world w, u realizes c at w and u realizes d at w, then c and d are identical; (.2) if, for some concept c and possible world w, u realizes c at w and v realizes c at w, then u is identical to v; and (.3) if, for some ordinary object u and concept c, u realizes c at w and u realizes c at w', then w is identical to w':

(.1) $\exists u \exists w(RealizesAt(u,c,w) \,\&\, RealizesAt(u,d,w)) \rightarrow c = d$

(.2) $\exists c \exists w(RealizesAt(u,c,w) \,\&\, RealizesAt(v,c,w)) \rightarrow u = v$

(.3) $\exists u \exists c(RealizesAt(u,c,w) \,\&\, RealizesAt(u,c,w')) \rightarrow w = w'$

(**700**) **Definition:** Appearance at a World. We say that a concept c *appears at* a possible world w just in case some ordinary object realizes c at w:

$$AppearsAt(c,w) \equiv_{df} \exists u RealizesAt(u,c,w)$$

(**701**) **Theorem:** Fact About Appearance. In light of the foregoing facts about realization at a world, we also have the following fact about appearance at a world: if a concept c appears at possible world w, then a unique ordinary object realizes c at w:

$$AppearsAt(c,w) \rightarrow \exists! u(RealizesAt(u,c,w))$$

(**702**) **Theorem:** Appearance and Being Ordinary. It proves useful to remember that if a concept c appears at a world w, then c encodes the property of *being ordinary*:

$$AppearsAt(c,w) \rightarrow cO!$$

(**703**) **Definition:** Mirroring. Recall that in (295) we stipulated that x encodes a proposition p, written $x\Sigma p$, just in case x exists and encodes $[\lambda y\, p]$. Since concepts are abstract objects (612), $c\Sigma p$ is defined and it follows that $c\Sigma p \equiv c[\lambda y\, p]$. So we now say that a concept c *mirrors* a possible world w just in case for any proposition p, c encodes p if and only if p is true at w:

$$Mirrors(c,w) \equiv_{df} \forall p(c\Sigma p \equiv w \models p)$$

Also, as noted in (515), $w \models p$ is equivalent to $w\Sigma p$. So the above definition entails the following equivalence:

$$Mirrors(c,w) \equiv \forall p(c\Sigma p \equiv w\Sigma p)$$

This better reveals why the definiens introduces a notion of mirroring: concept c mirrors a possible world w just in case c and w encode the same propositions.

(**704**) **Theorem:** Appearance and Mirroring. It now follows that if a concept appears at a possible world, it mirrors that world:

$$AppearsAt(c,w) \rightarrow Mirrors(c,w)$$

It is important to understand why the right-to-left direction fails, i.e., why it is provable that $\exists c \exists w(Mirrors(c,w)\ \&\ \neg AppearsAt(c,w))$. To find witnesses to this claim, consider any possible world, say w_1 (we know that there are are at least two, by (547.4)). Since possible worlds are abstract objects, they are also concepts. So w_1 is both a possible world and a concept. Then, clearly, $Mirrors(w_1,w_1)$. However, $\neg AppearsAt(w_1,w_1)$. For suppose $AppearsAt(w_1,w_1)$, for reductio. Then, by our useful fact (702), it follows that $w_1 O!$. But since w_1 is, by hypothesis, a possible world, it is a situation (512). Hence every property w_1 encodes is a propositional property (467). It follows that $Propositional(O!)$. But this contradicts (279.4.c).

(**705**) **Theorem:** New Fact About Appearance. If some concept appears at possible worlds w and w', then $w = w'$:

$$\exists c(AppearsAt(c,w)\ \&\ AppearsAt(c,w')) \rightarrow w = w'$$

(**706**) **Theorem:** Appearance At is Rigid. A concept c appears at a possible world w if and only if it necessarily does so:

$$AppearsAt(c,w) \equiv \Box AppearsAt(c,w)$$

(**707**) **Theorem:** New Fact About Realization. We can now easily prove that if some concept c is such that u realizes c at w and v realizes c at w', then both w is identical to w' and u is identical to v.

$$\exists c(RealizesAt(u,c,w)\ \&\ RealizesAt(v,c,w')) \rightarrow (w = w'\ \&\ u = v)$$

Intuitively, this theorem tells us that the concept alone fixes the other parameters of the condition $RealizesAt(u,c,w)$.

(**708**) **Lemma:** Concepts of Ordinary Individuals, Realization, Appearance, and Mirroring.[346] (.1) u realizes the concept of u at the actual world; (.2) the concept of u appears at the actual world; (.3) the concept of u mirrors the actual world:

[346] I'm indebted to Daniel West for pointing out that these theorems can be proved by modally strict means.

(.1) $RealizesAt(u, \boldsymbol{c}_u, \boldsymbol{w}_\alpha)$

(.2) $AppearsAt(\boldsymbol{c}_u, \boldsymbol{w}_\alpha)$

(.3) $Mirrors(\boldsymbol{c}_u, \boldsymbol{w}_\alpha)$

13.4.2 Possible-Individual Concepts

(**709**) **Definition:** Possible-Individual Concepts. We've previously introduced (a) the notion of a concept of an ordinary individual, i.e., $ConceptOf(c,u)$ (680, and (b) the concept of individual u, i.e., c_u (682). These are notions that were defined without reference to possible worlds. We now introduce the notion of a *possible-individual concept* as any concept that appears at some possible world:

$$PossibleIndividualConcept(c) \equiv_{df} \exists w AppearsAt(c,w)$$

It should be clear that by 'possible-individual' concept, we mean a concept of an ordinary individual. One could broaden the notion of a possible-individual concept by reinterpreting the variable u in this Chapter so that it ranges over discernibles. But for the reasons given in (679), we'll operate with the narrower understanding defined above.

(**710**) **Theorem:** Concepts of Ordinary Individuals are Possible-Individual Concepts. It now follows that if c is a concept of some ordinary individual u, then c is a possible-individual concept:

$$\exists u ConceptOf(c,u) \rightarrow PossibleIndividualConcept(c)$$

Note that the converse fails. As an exercise, the reader should prove that there exists a possible-individual concept c that is not a concept of any individual u, i.e., prove $\exists c(PossibleIndividualConcept(c)\ \&\ \neg\exists u ConceptOf(c,u))$.[347]

There is an easy, but non-modally strict proof that appeals to theorem (536)★, i.e., $p \equiv \boldsymbol{w}_\alpha \models p$.[348] Yet with some work, this theorem can be proved by modally strict means.

(**711**) **Theorem:** The Concept of u is a Possible-Individual Concept.

[347] In previous work (Zalta 2000, §8.2), I used 'IndividualConcept' instead of 'PossibleIndividualConcept' for the notion defined in (709). I'm indebted to Daniel West for pointing out that, if we had preserved this usage from the 2000 paper, the claim just left as an exercise converse could plausibly be read as: there is an individual concept that is not the concept of any individual. So I've revised the definiendum in (709) to avoid this reading.

[348] Assume $\exists u ConceptOf(c,u)$, and suppose a be such an ordinary object, so that we know $ConceptOf(c,a)$. If we can show $\forall F(\boldsymbol{w}_\alpha \models Fa \equiv cF)$, then by existentially generalizing on a and $\boldsymbol{w}_\alpha$ and applying definitions (697), (700), and (709), we're done. But by definition (680), $ConceptOf(c,a)$ implies $\forall F(cF \equiv Fa)$. Moreover, as an instance of (536)★, we know $Fa \equiv \boldsymbol{w}_\alpha \models Fa$, which by GEN yields $\forall F(Fa \equiv \boldsymbol{w}_\alpha \models Fa)$. So by the laws of quantified biconditionals, $\forall F(cF \equiv \boldsymbol{w}_\alpha \models Fa)$, i.e., $\forall F(\boldsymbol{w}_\alpha \models Fa \equiv cF)$.

$$IndividualConcept(\boldsymbol{c}_u)$$

(**712**) **Theorem:** Rigidity of *PossibleIndividualConcept* and Restricted Variables. By our conventions for restricted variables, definition (709) abbreviates $PossibleIndividualConcept(x) \equiv_{df} C!x \,\&\, \exists y(PossibleWorld(y) \,\&\, AppearsAt(x,y))$. It should be clear that $IndividualConcept(x)$ is a restriction condition (336). It also follows, however, that $PossibleIndividualConcept(x)$ is a rigid restriction condition (337), since it is a modally strict theorem that every possible-individual concept is necessarily a possible-individual concept:

$$\forall x(PossibleIndividualConcept(x) \to \Box PossibleIndividualConcept(x))$$

We may therefore introduce the circumflexed, lower-case italic letters $\hat{c}, \hat{d}, \hat{e}, \ldots$ as rigid restricted variables ranging over possible-individual concepts. Note that the present theorem guarantees that the quantifiers $\forall\hat{c}$ and $\exists\hat{c}$ behave classically in the sense that $\forall\hat{c}\varphi \to \exists\hat{c}\varphi$; cf. Remark (342). Note also that the remarks about doubly restricted variables in (514) apply to our restricted variables for individual concepts. For example, we have two options for expanding $\forall\hat{c}\varphi^{\hat{c}}_x$, namely, either as $\forall x(PossibleIndividualConcept(x) \to \varphi)$ or as $\forall c(PossibleIndividualConcept(c) \to \varphi^c_x)$.

(**713**) **Theorem:** Appearance at a Unique Possible World. If we add parenthetical quantifiers to improve readability, we may now establish (.1) (for any possible-individual concept $\hat{c}$), there is a unique possible world at which $\hat{c}$ appears; and (.2) (for any possible-individual concept $\hat{c}$), there is a unique world w such that necessarily, $\hat{c}$ appears at w:

(.1) $\exists!w AppearsAt(\hat{c}, w)$

(.2) $\exists!w \Box AppearsAt(\hat{c}, w)$

By applying the Rule of Actualization to (.1) and appealing to theorem (176.2), it follows that (.3) the world at which $\hat{c}$ appears exists:

(.3) $\imath w AppearsAt(\hat{c}, w)\downarrow$

Exercise: Give a proof of (.3) without appealing to the Rule of Actualization or theorem (176.2), but instead uses theorem (174.3) and recent theorems.

(**714**) **Definition:** The Possible World At Which a Possible-Individual Concept Appears. The previous theorem allows us to introduce the notation $\boldsymbol{w}_{\hat{c}}$ to refer to the possible world at which possible-individual concept $\hat{c}$ appears:

$$\boldsymbol{w}_{\hat{c}} =_{df} \imath w AppearsAt(\hat{c}, w)$$

(**715**) **Theorem:** Modally Strict Facts About $\boldsymbol{w}_{\hat{c}}$. It is now a modally strict fact that (.1) a possible-individual concept appears at the world where it appears:

(.1) $AppearsAt(\hat{c}, \boldsymbol{w}_{\hat{c}})$

This theorem sounds trivial and there is a non-modally strict proof that *is* trivial, which appeals to (145.2)★. But the modally strict proof involves (713.2), theorem (153.1), and definition (714).

It thus follows that (.2) a possible-individual concept $\hat{c}$ mirrors the possible world where it appears:

(.2) $Mirrors(\hat{c}, \boldsymbol{w}_{\hat{c}})$

Cf. Leibniz 1714, Section 56, where we find (PW 187, G.vi 616):

> Now this *connexion* or adaptation of all created things with each, and of each with all the rest, means that each simple substance has relations which express all the others, and that consequently, it is a perpetual living mirror of the universe.

Here, we have to interpret Leibniz's talk of simple substances in terms of the individual concepts of ordinary individuals. Of course, it is not clear what to make of Leibniz's suggestion that a simple substance is a *living* mirror.

In any case, previous theorems and definitions now yield that (.3) a possible-individual concept $\hat{c}$ contains the possible world where it appears:

(.3) $\hat{c} \succeq \boldsymbol{w}_{\hat{c}}$

Cf. Leibniz 1686, Article 9, where we find (PW 19–20, G.iv 434):

> Further, every substance is like an entire world and like a mirror of God, or of the whole universe, which each one expresses in its own way, very much as one and the same town is variously represented in accordance with different positions of the observer. Thus, the universe is in a way multiplied as many times as there are substances, ...

So the metaphor of mirroring was used early on in Leibniz's work. Once we interpret Leibniz to be talking about the possible-individual concepts of ordinary individuals, we see that the world where a possible-individual concept appears ($\boldsymbol{w}_{\hat{c}}$) is indeed 'multiplied as many times as there are substances' since each $\hat{c}$ that appears at $\boldsymbol{w}_{\hat{c}}$ has $\boldsymbol{w}_{\hat{c}}$ as a part.

(**716**) **Theorem:** Possible-Individual Concepts Contain the Concepts of Encoded Properties. It is also a consequence of the foregoing that $\hat{c}$ encodes a property if and only if it contains the concept of that property:

$$\hat{c}G \equiv \hat{c} \succeq \boldsymbol{c}_G$$

(**717**) **Theorems:** Possible-Individual Concepts and Property Negation. Facts about property negation become reflected in individual concepts as follows:

(.1) a possible-individual concept encodes G if and only if it fails to encode the negation of G; (.2) a possible-individual concept encodes the negation of G if and only if it fails to encode G; (.3) a possible-individual concept contains the concept of G if and only if it doesn't contain the concept of the negation of G; and (.4) a possible-individual concept doesn't contain the concept of G if and only if it contains the concept of the negation of G:

(.1) $\hat{c}G \equiv \neg\hat{c}\overline{G}$

(.2) $\hat{c}\overline{G} \equiv \neg\hat{c}G$

(.3) $\hat{c} \succeq \boldsymbol{c}_G \equiv \hat{c} \not\succeq \boldsymbol{c}_{\overline{G}}$

(.4) $\hat{c} \not\succeq \boldsymbol{c}_G \equiv \hat{c} \succeq \boldsymbol{c}_{\overline{G}}$

(**718**) **Theorem:** Possible-Individual Concepts and Completeness. Recall that in (687) we defined a sense in which concepts are complete. It straightforwardly follows that a possible-individual concept is complete:

$$Complete(\hat{c})$$

This provides further confirmation of Article 8 of the *Discourse on Metaphysics*, quoted earlier, where Leibniz says that "it is in the nature of an individual substance, or complete being, to have a notion so complete that it is sufficient to contain and render deducible from itself, all the predicates of the subject to which this notion is attributed" (PW 18–19, G.iv 433).

13.4.3 Compossibility

(**719**) **Definition:** Compossibility. We say that two possible-individual concepts are *compossible* just in case they appear at the same possible world:

$$Compossible(\hat{c},\hat{e}) \equiv_{df} \exists w(AppearsAt(\hat{c},w) \,\&\, AppearsAt(\hat{e},w))$$

(**720**) **Lemma:** A Common Possible World. It follows from the previous definition that possible-individual concepts $\hat{c}$ and $\hat{e}$ are compossible if and only if the possible world where $\hat{c}$ appears is identical to the one where $\hat{e}$ appears:

$$Compossible(\hat{c},\hat{e}) \equiv \boldsymbol{w}_{\hat{c}} = \boldsymbol{w}_{\hat{e}}$$

(**721**) **Theorems:** Compossibility is an Equivalence Condition on Possible-Individual Concepts. It follows from the previous lemma that compossibility is a reflexive, symmetric, and transitive condition with respect to possible-individual concepts:

(.1) $Compossible(\hat{c},\hat{c})$

(.2) $Compossible(\hat{c},\hat{e}) \rightarrow Compossible(\hat{e},\hat{c})$

(.3) $Compossible(\hat{c},\hat{d}) \,\&\, Compossible(\hat{d},\hat{e}) \rightarrow Compossible(\hat{c},\hat{e})$

Since compossibility is an equivalence condition with respect to possible-individual concepts, we know that the latter are partitioned. In light of lemma (720), all of the possible-individual concepts that are compossible with one another appear at a common possible world.

13.4.4 Counterparts

(722) **Definition:** Counterpart Of. We say that $\hat{e}$ is a *counterpart of* $\hat{c}$ just in case there is an ordinary individual u and there are possible worlds w and w' such that u realizes $\hat{c}$ at w and u realizes $\hat{e}$ at w':

$$CounterpartOf(\hat{e},\hat{c}) \equiv_{df} \exists u \exists w \exists w'(RealizesAt(u,\hat{c},w) \,\&\, RealizesAt(u,\hat{e},w'))$$

For example, if Alexander is an ordinary object that realizes an individual concept $\hat{c}$ at the actual world and realizes an individual concept $\hat{e}$ at some non-actual possible world, then $\hat{e}$ is a counterpart of $\hat{c}$.

(723) **Theorem:** Counterpart Of is an Equivalence Condition on Possible-Individual Concepts. We now have (.1) a possible-individual concept is a counterpart of itself; (.2) if $\hat{e}$ is a counterpart of $\hat{c}$, then $\hat{c}$ is a counterpart of $\hat{e}$; and (.3) if $\hat{e}$ is a counterpart of $\hat{d}$, and $\hat{d}$ is a counterpart of $\hat{c}$, then $\hat{e}$ is a counterpart of $\hat{c}$:

(.1) $CounterpartOf(\hat{c},\hat{c})$

(.2) $CounterpartOf(\hat{e},\hat{c}) \rightarrow CounterpartOf(\hat{c},\hat{e})$

(.3) $CounterpartsOf(\hat{e},\hat{d}) \,\&\, CounterpartOf(\hat{d},\hat{c}) \rightarrow CounterpartOf(\hat{e},\hat{c})$

(724) **Theorem:** Counterparts and Realization. It follows that if $\hat{e}$ is a counterpart of $\hat{c}$, then there is a unique ordinary individual that realizes $\hat{c}$ at some world w and that realizes $\hat{e}$ at some world w':

$$CounterpartOf(\hat{e},\hat{c}) \equiv \exists! u \exists w \exists w'(RealizesAt(u,\hat{c},w) \,\&\, RealizesAt(u,\hat{e},w'))$$

13.4.5 World-Relative Concepts of Individuals

(725) **Definition:** World-Relative Concepts of Ordinary Individuals. Let us say that c is a *concept of u at w* just in case c encodes exactly the properties that u exemplifies at w:

$$ConceptOfAt(c,u,w) \equiv_{df} \forall F(cF \equiv w \models Fu)$$

Note that we could have used $RealizesAt(u,c,w)$ as the definiens for the definiendum $ConceptOfAt(c,u,w)$, since the definiens of $RealizesAt(u,c,w)$ is equivalent to the one above; they differ only by commuting the (quantified) biconditional. Though the linguistic notion of *focus* doesn't apply to the formulas in our system, it nevertheless has some application to the regimented natural language we use to read those formulas. When we read the above definiens, the focus is on the concept c and the properties it encodes, whereas when we read the definiens of $RealizesAt(u,c,w)$, the focus is on u and the properties it exemplifies at w.

(**726**) **Theorems:** Existence of World-Relative Concepts of Individuals. It follows that (.1) there is a concept of u at w; (.2) there is a unique concept of u at w; and (.3) the concept of u at w exists:

(.1) $\exists c ConceptOfAt(c,u,w)$

(.2) $\exists! c ConceptOfAt(c,u,w)$

(.3) $\imath c ConceptOfAt(c,u,w)\downarrow$

Strictly speaking, these are conditional existence claims, given the presence of the restricted variables u and w. But since we know that ordinary individuals exist, by (227.1) and the T schema, and that there are at least two possible worlds (547.4), we can derive *unconditional* existence claims.

(**727**) **Definition:** Notation for the Concept of u at w. We henceforth use the notation $\boldsymbol{c}_u^w$ to denote the concept of u at w:

$$\boldsymbol{c}_u^w =_{df} \imath c ConceptOfAt(c,u,w)$$

This introduces $\boldsymbol{c}_u^w$ as a binary functional term with the free restricted variables u and w.

(**728**) **Theorems:** $\boldsymbol{c}_u^w$ is Strictly Canonical. Clearly, it follows that (.1) $\boldsymbol{c}_u^w$ is (identical to) a canonical concept:

(.1) $\boldsymbol{c}_u^w = \imath c \forall F(cF \equiv w \models Fu)$

If we let φ be the formula $w \models Fu$, then it follows that φ is a rigid condition on properties, i.e., that (.2) necessarily, every property such that φ is necessarily such that φ:

(.2) $\Box \forall F(w \models Fu \to \Box w \models Fu)$

So $\boldsymbol{c}_u^w$ is (identical to) a strictly canonical concept, by (260.2). Hence, by (261.2), it follows that (.3) $\boldsymbol{c}_u^w$ is a concept that encodes exactly those properties F that u exemplifies at w:

(.3) $C!\boldsymbol{c}_u^w \mathbin{\&} \forall F(\boldsymbol{c}_u^w F \equiv w \models Fu)$

(**729**) **Lemma:** Basic Facts About World-Relative Concepts of Ordinary Individuals. (.1) u realizes the-concept-of-u-at-w at w; (.2) the concept of u at w appears at w; (.3) the concept of u at w is an individual concept; (.4) the concept of u at w mirrors w; and (.5) the concept of u at w is complete:

(.1) $\mathit{RealizesAt}(u, \boldsymbol{c}_u^w, w)$

(.2) $\mathit{AppearsAt}(\boldsymbol{c}_u^w, w)$

(.3) $\mathit{PossibleIndividualConcept}(\boldsymbol{c}_u^w)$

(.4) $\mathit{Mirrors}(\boldsymbol{c}_u^w, w)$

(.5) $\mathit{Complete}(\boldsymbol{c}_u^w)$

These are all modally strict.

(**730**) **Theorem:** Equivalence of Possible-Individual Concepts and World-Relative Concepts of Individuals. From the definitions of possible-individual concept and world-relative concept of an individual, we may prove that (.1) a concept c is a possible-individual concept iff there is some ordinary object u and possible world w such that c is a concept of u at w; and (.2) a concept c is a possible-individual concept iff there is some and ordinary object u and possible world w such that c is identical to the concept of u at w:

(.1) $\mathit{PossibleIndividualConcept}(c) \equiv \exists u \exists w \mathit{ConceptOfAt}(c, u, w)$

(.2) $\mathit{PossibleIndividualConcept}(c) \equiv \exists u \exists w (c = \boldsymbol{c}_u^w)$

(**731**) **Lemma:** The Concept of an Individual at the Actual World. It is a basic fact about the notions that we've defined so far that the concept of u at the actual world $\boldsymbol{w}_\alpha$ is identical to the concept of u:

$$\boldsymbol{c}_u^{\boldsymbol{w}_\alpha} = \boldsymbol{c}_u$$

While it is easy to give a non-modally strict proof of this theorem, the interest is in the fact that it is capable of a modally strict proof.

(**732**) **Lemmas:** Further Facts About World-Relative Concepts of Individuals. The following 5 lemmas are also immediately forthcoming: (.1) the concept of u at w encodes a property G iff it contains $\boldsymbol{c}_G$; (.2) u exemplifies G at w iff the concept of u at w contains $\boldsymbol{c}_G$; (.3) if the concept of u at w is identical to the concept of v at w, then u and v are identical; (.4) if the concept of u at w is identical to the concept of u at w', then $w = w'$; and (.5) if the concept of u at w is identical to the concept of v at w', then $w = w'$ and $u = v$:

(.1) $\boldsymbol{c}_u^w G \equiv \boldsymbol{c}_u^w \succeq \boldsymbol{c}_G$

(.2) $w \models Gu \equiv \boldsymbol{c}_u^w \succeq \boldsymbol{c}_G$

(.3) $\boldsymbol{c}_u^w = \boldsymbol{c}_v^w \rightarrow u = v$

(.4) $\boldsymbol{c}_u^w = \boldsymbol{c}_u^{w'} \rightarrow w = w'$

(.5) $\boldsymbol{c}_u^w = \boldsymbol{c}_v^{w'} \rightarrow (w = w' \;\&\; u = v)$

(**733**) **Theorem:** Compossibility, Counterparts, and World-Relative Concepts of Individuals. Some of the most important facts about world-relative concepts of individuals are: (.1) the concept of u at w and the concept of v at w are compossible; (.2) $\hat{c}$ and $\hat{e}$ are compossible if and only if there are an ordinary individuals u and v and a possible world w such that $\hat{c}$ is the concept of u at w and $\hat{e}$ is the concept of v at w; (.3) the-concept-of-u-at-w' is a counterpart of the-concept-of-u-at-w; and (.4) $\hat{e}$ is a counterpart of $\hat{c}$ if and only if there is an ordinary individual u and there are worlds w and w' such that $\hat{c}$ is the concept of u at w and $\hat{e}$ is the concept of u at w':

(.1) $\mathit{Compossible}(\boldsymbol{c}_u^w, \boldsymbol{c}_v^w)$

(.2) $\mathit{Compossible}(\hat{c}, \hat{e}) \equiv \exists u \exists v \exists w(\hat{c} = \boldsymbol{c}_u^w \;\&\; \hat{e} = \boldsymbol{c}_v^w)$

(.3) $\mathit{CounterpartOf}(\boldsymbol{c}_u^{w'}, \boldsymbol{c}_u^w)$

(.4) $\mathit{CounterpartOf}(\hat{e}, \hat{c}) \equiv \exists u \exists w \exists w'(\hat{c} = \boldsymbol{c}_u^w \;\&\; \hat{e} = \boldsymbol{c}_u^{w'})$

When Leibniz speaks of 'the several Sextuses' and 'the many possible Adams', we take him to be referring to the world-relatized concepts of these individuals. The world-relativized concepts of Sextus are counterparts, as are the world-relativized concepts of Adam. It will soon become apparent that (.3) plays an important role in a fundamental theorem of Leibnizian modal metaphysics.

(**734**) **Remark:** Compossibility and Leibniz's Theodicy. Though the theory of individual concepts and possible worlds articulated thus far is completely silent about the existence of God, it nevertheless preserves an element of Leibniz's theodicy, namely, his conception of the work that God would have had to undertake in order to 'create' the actual world. Leibniz makes it clear that God had to first evaluate all the possible worlds to determine which one should be actualized. But our works shows that to evaluate the possible worlds, all God had to do was to inspect an arbitrarily chosen possible-individual concept from each group of compossible such concepts. That one inspection alone reveals all the facts about *the* world where that possible-individual concept appears, since every possible-individual concept of the group mirrors that world.

Note that our metaphysics identifies the actual world only as the possible world that encodes, in the sense of (295), all and only the true propositions. So, having decided that some particular possible world w was the best, God could 'actualize' w by making it the case that every proposition encoded in w is true. This, in effect, creates a physical universe, as defined in (546). Indeed, when inspecting the possible worlds by examining an arbitrary possible-individual concept from each group of compossible such concepts, God could have just 'actualized' the possible-individual concept $\hat{c}$ that was chosen for the determination of whether $w_{\hat{c}}$ was the best. To do so, God would have to make it the case that there exists an ordinary object which in fact exemplifies all the properties that $\hat{c}$ encodes. In doing *that*, God would as a consequence actualize $\boldsymbol{w}_{\hat{c}}$, for when God made it the case that there is an ordinary object that exemplifies the properties $\hat{c}$ encodes, God would have also made it the case that all of the propositions true in $\boldsymbol{w}_{\hat{c}}$ are true, since $\hat{c}$ mirrors $\boldsymbol{w}_{\hat{c}}$ (715.2).

(**735**) **Remark:** Another Reason Why Concepts Are Not Properties. It is worth mentioning here that another reason not to identify concepts as properties is that such a view gets the Leibnizian *metaphysics* of individual concepts wrong. It is central to Leibniz's view of individual concepts that for each ordinary individual u and possible world w, a unique possible-individual concept corresponds to u at w. But if concepts are analyzed as properties, and possible-individual concepts become a special kind of property, then uniqueness will fail unless the property theorist *requires* that properties be identical when necessarily equivalent, i.e., that $\Box\forall x(Fx \equiv Gx) \rightarrow F = G$.

To see why, suppose one were to define:

> F is a possible-individual concept of u at w if and only if both (a) F necessarily implies all of the properties that u exemplifies at w, and (b) u uniquely exemplifies F at w.

To see the problem with this definition, consider the concept of Adam. Since the necessary equivalence of properties doesn't imply their identity, one can't prove that for every world w, there is a unique possible-individual concept of Adam at w, as defined above. For suppose property P is the possible-individual concept of Adam at w. Now suppose Q is a property necessarily equivalent to, but distinct from, P. Then Q is an individual concept of Adam, by the following reasoning:

- Q satisfies clause (a) of the definition. Since P and Q are necessarily equivalent, then they necessarily imply the same properties by (443.3). So if P necessarily implies all the properties Adam exemplifies at w, then so does Q.

- Q satisfies clause (b) of the definition. Since Adam uniquely exemplifies

P at w, and P and Q are necessarily equivalent, then Adam uniquely exemplifies Q at w (we leave the proof as an exercise).

Since necessary equivalence doesn't imply identity, and there are properties like Q that are necessarily equivalent to P, then P isn't a unique possible-individual concept of Adam at w. Thus, the property theorist faces a dilemma: either require that necessarily equivalent properties are identical and derive the existence of a unique concept of Adam at w or omit the requirement that necessarily equivalent properties are identical and give up the claim that there is a unique concept of Adam at w. Such a dilemma is not faced on the present analysis.

13.4.6 Fundamental Theorems

(**736**) ⋆**Theorem:** A Fundamental Theorem of Leibnizian Modal Metaphysics. It seems reasonable to suggest that Leibniz's modal metaphysics is driven by the following conditionals (even though Leibniz never explicitly formulates them as such), namely, (.1) if an ordinary individual u exemplifies F but might not have, then both (a) the concept of u contains the concept of F and (b) there is a possible-individual concept $\hat{c}$ such that $\hat{c}$ is a counterpart of the concept of u, $\hat{c}$ doesn't contain the concept of F, and $\hat{c}$ appears at a possible world distinct from the actual world; and (.2) if an ordinary individual u doesn't exemplify F but might have, then both (a) the concept of u doesn't contain the concept of F and (b) there is a possible-individual concept $\hat{c}$ such that $\hat{c}$ is a counterpart of the concept of u, $\hat{c}$ does contain the concept of F, and $\hat{c}$ appears at a possible world distinct from the actual world:

(.1) $Fu \,\&\, \Diamond\neg Fu \rightarrow$
$\boldsymbol{c}_u \succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(\mathit{CounterpartOf}(\hat{c},\boldsymbol{c}_u) \,\&\, \hat{c} \not\succeq \boldsymbol{c}_F \,\&\, \exists w(w \neq \boldsymbol{w}_\alpha \,\&\, \mathit{AppearsAt}(\hat{c},w)))$

(.2) $\neg Fu \,\&\, \Diamond Fu \rightarrow$
$\boldsymbol{c}_u \not\succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(\mathit{CounterpartOf}(\hat{c},\boldsymbol{c}_u) \,\&\, \hat{c} \succeq \boldsymbol{c}_F \,\&\, \exists w(w \neq \boldsymbol{w}_\alpha \,\&\, \mathit{AppearsAt}(\hat{c},w)))$

If we suppose our theory has been applied, then as an example of (.1), we have: if Alexander is king but might not have been, then the concept of Alexander contains the concept of *being a king* and there is a possible-individual concept that is a counterpart of the concept of Alexander, that doesn't contain the concept of *being a king*, and that appears at some world other than the actual world. As an example of (.2), we have: if Alexander fails to be a philosopher but might have been, then the concept of Alexander fails to contain the concept of *being a philosopher* and there is a possible-individual concept that is a counterpart of the concept of Alexander, that does contain the concept of *being a philosopher*, and that appears at some world other than the actual world.

(**737**) ⋆**Theorems:** Biconditional Fundamental Theorems. Though we've labeled the preceding theorem a fundamental theorem of Leibnizian modal metaphysics, it is reasonable to suggest that, strictly speaking, a fundamental theorem so-called should be a biconditional. Though the converses of the preceding theorems are indeed derivable, a study of the matter reveals that these converses have antecedents that are stronger than they need to be. We shall see, for example, that one can derive $Fu\ \&\ \Diamond\neg Fu$ from the simpler conjunction $\boldsymbol{c}_u \succeq \boldsymbol{c}_F\ \&\ \exists\hat{c}(CounterpartOf(\hat{c},\boldsymbol{c}_u)\ \&\ \hat{c} \not\succeq \boldsymbol{c}_F)$. This shows how much information is packed into the notions of concept containment and counterparts.

Consequently, we can formulate biconditional fundamental theorems as follows: (.1) an ordinary individual u exemplifies F but might not have if and only if both (a) the concept of u contains the concept of F and (b) there is a possible-individual concept $\hat{c}$ such that $\hat{c}$ is a counterpart of the concept of u and $\hat{c}$ doesn't contain the concept of F; and (.2) an ordinary individual u doesn't exemplify F but might have if and only if both (a) the concept of u doesn't contain the concept of F and (b) there is a possible-individual concept $\hat{c}$ such that $\hat{c}$ is a counterpart of the concept of u and $\hat{c}$ does contain the concept of F:

(.1) $Fu\ \&\ \Diamond\neg Fu \equiv \boldsymbol{c}_u \succeq \boldsymbol{c}_F\ \&\ \exists\hat{c}(CounterpartOf(\hat{c},\boldsymbol{c}_u)\ \&\ \hat{c} \not\succeq \boldsymbol{c}_F)$

(.2) $\neg Fu\ \&\ \Diamond Fu \equiv \boldsymbol{c}_u \not\succeq \boldsymbol{c}_F\ \&\ \exists\hat{c}(CounterpartOf(\hat{c},\boldsymbol{c}_u)\ \&\ \hat{c} \succeq \boldsymbol{c}_F)$

These theorems constitute *Leibnizian truth conditions*, respectively, for the claim that Fu is contingently true and the claim that Fu is contingently false.

(**738**) **Theorems:** Related, Modally Strict Theorems. It is interesting to note that there are several ways of adjusting the formulation of the fundamental theorems so as to produce modally strict versions. The most basic of these ways is to relativize both sides of the conditionals in (736)⋆, and both sides of the biconditionals in (737)⋆, to a possible world. This yields the facts that (.1) if it is true at possible world w that an ordinary individual u exemplifies F but might not have, then (a) the concept of u at w contains the concept of F and (b) there is a possible-individual concept $\hat{c}$ such that $\hat{c}$ is a counterpart of the concept of u-at-w, $\hat{c}$ doesn't contain the concept of F, and $\hat{c}$ appears at a possible world distinct from w; and (.2) if it is true, at possible world w, that an ordinary individual u doesn't exemplify F but might have, then both (a) the concept of u at w doesn't contain the concept of F and (b) there is a possible-individual concept $\hat{c}$ such that $\hat{c}$ is a counterpart of the concept of u-at-w, $\hat{c}$ does contain the concept of F, and $\hat{c}$ appears at a possible world distinct from w; (.3) it is true, at possible world w, that an ordinary individual u exemplifies F but might not have if and only if (a) the concept of u at w contains the concept of F and (b) there is a possible-individual concept $\hat{c}$ such that $\hat{c}$ is a counterpart of the concept of u-at-w and $\hat{c}$ doesn't contain the concept of F; and (.4) it is true, at possible world w, that an ordinary individual u doesn't exemplify F

but might have if and only if both (a) the concept of u at w doesn't contain the concept of F and (b) there is a possible-individual concept $\hat{c}$ such that $\hat{c}$ is a counterpart of the concept of u-at-w and $\hat{c}$ does contain the concept of F:

(.1) $w \models (Fu \,\&\, \Diamond\neg Fu) \rightarrow$
$\boldsymbol{c}_u^w \succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^w) \,\&\, \hat{c} \not\succeq \boldsymbol{c}_F \,\&\, \exists w'(w' \neq w \,\&\, AppearsAt(\hat{c}, w')))$

(.2) $w \models (\neg Fu \,\&\, \Diamond Fu) \rightarrow$
$\boldsymbol{c}_u^w \not\succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^w) \,\&\, \hat{c} \succeq \boldsymbol{c}_F \,\&\, \exists w'(w' \neq w \,\&\, AppearsAt(\hat{c}, w')))$

(.3) $w \models (Fu \,\&\, \Diamond\neg Fu) \equiv \boldsymbol{c}_u^w \succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^w) \,\&\, \hat{c} \not\succeq \boldsymbol{c}_F)$

(.4) $w \models (\neg Fu \,\&\, \Diamond Fu) \equiv \boldsymbol{c}_u^w \not\succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^w) \,\&\, \hat{c} \succeq \boldsymbol{c}_F)$

Of course, given our discussion of Leibniz's notion of hypothetical necessity, it may be that Leibniz would not have felt the need for modally strict versions of the fundamental theorems.

Exercises: Show that the following versions of (736.1)★ and (736.2)★ are modally strict:

(.1) $\Diamond(Fu \,\&\, \Diamond\neg Fu) \rightarrow \exists w(\boldsymbol{c}_u^w \succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^w) \,\&\, \hat{c} \not\succeq \boldsymbol{c}_F \,\&$
$\exists w'(w' \neq w \,\&\, AppearsAt(\hat{c}, w'))))$

(.2) $\Diamond(\neg Fu \,\&\, \Diamond Fu) \rightarrow \exists w(\boldsymbol{c}_u^w \not\succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^w) \,\&\, \hat{c} \succeq \boldsymbol{c}_F \,\&$
$\exists w'(w' \neq w \,\&\, AppearsAt(\hat{c}, w'))))$

Also, show that the following versions of (736.1)★ and (736.2)★ are modally strict:

(.3) $\mathcal{A}(Fu \,\&\, \Diamond\neg Fu) \rightarrow \boldsymbol{c}_u^{w_\alpha} \succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^{w_\alpha}) \,\&\, \hat{c} \not\succeq \boldsymbol{c}_F \,\&$
$\exists w'(w' \neq w_\alpha \,\&\, AppearsAt(\hat{c}, w')))$

(.4) $\mathcal{A}(\neg Fu \,\&\, \Diamond Fu) \rightarrow \boldsymbol{c}_u^{w_\alpha} \not\succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^{w_\alpha}) \,\&\, \hat{c} \succeq \boldsymbol{c}_F \,\&$
$\exists w'(w' \neq w_\alpha \,\&\, AppearsAt(\hat{c}, w')))$

Now show that the following versions of (737.1)★ and (737.2)★ are modally strict:

(.5) $\Diamond(Fu \,\&\, \Diamond\neg Fu) \equiv \exists w(\boldsymbol{c}_u^w \succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^w) \,\&\, \hat{c} \not\succeq \boldsymbol{c}_F))$

(.6) $\Diamond(\neg Fu \,\&\, \Diamond Fu) \equiv \exists w(\boldsymbol{c}_u^w \not\succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^w) \,\&\, \hat{c} \succeq \boldsymbol{c}_F))$

Also, show that the following versions of (737.1)★ and (737.2)★ are modally strict:

(.7) $\mathcal{A}(Fu \,\&\, \Diamond\neg Fu) \equiv \boldsymbol{c}_u^{\boldsymbol{w}_\alpha} \succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^{\boldsymbol{w}_\alpha}) \,\&\, \hat{c} \not\succeq \boldsymbol{c}_F)$

(.8) $\mathcal{A}(\neg Fu \,\&\, \Diamond Fu) \equiv \boldsymbol{c}_u^{\boldsymbol{w}_\alpha} \not\succeq \boldsymbol{c}_F \,\&\, \exists\hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^{\boldsymbol{w}_\alpha}) \,\&\, \hat{c} \succeq \boldsymbol{c}_F)$

(**739**) **Remark:** Recall the skeptical passage in Stalnaker 1976 (65) quoted in Remark (510):

> According to Leibniz, the universe—the actual world—is one of an infinite number of possible worlds existing in the mind of God. God created the universe by actualizing one of these possible worlds—the best one. It is a striking image, this picture of an infinite swarm of total universes, each by its natural inclination for existence striving for a position that can be occupied by only one, with God, in his infinite wisdom and benevolence, settling the competition by selecting the most worthy candidate. But in these enlightened times, we find it difficult to take this metaphysical myth any more seriously than the other less abstract creation stories told by our primitive ancestors. Even the more recent expurgated versions of the story, leaving out God and the notoriously chauvinistic thesis that our world is better than all the rest, are generally regarded, at best, as fanciful metaphors for a more sober reality.

By all means, let us leave out God and the chauvinistic thesis that our world is better than all the rest. I hope the foregoing effort at reconstructing Leibniz's views shows that they were not mere fanciful metaphors. The importance of the theorems in this final section shouldn't be understated. An unanalyzed truth to the effect that an ordinary object exemplifies a property but might not have (or that an ordinary object doesn't exemplify a property but might have), which is represented and regimented in terms of our modern notions of exemplification (Fx), negation ($\neg$) and possibility ($\Diamond$), implies a complex web of facts in Leibniz's modal metaphysics, involving the notions of: concepts ($C!$), concept containment ($\succeq$), possible worlds (w), identity ($=$), concepts of individuals ($\boldsymbol{c}_u$), possible-individual concepts ($\hat{c}$), concepts of properties ($\boldsymbol{c}_F$), appearance, and counterpart of. And this list doesn't include the notions in terms of which these notions are defined, such as encoding (xF), abstract objects ($A!$), situations, truth in a situation ($\models$), and propositional properties ($[\lambda y\, p]$).

Moreover, these theorems have powerful consequences when the theory is applied. Once we start extending our theory with familiar, uncontroversial truths about the properties that ordinary objects exemplify (or fail to exemplify) contingently, an elaborate network of truths involving primitive and defined notions emerges and describes an elegant and precise metaphysical picture that articulates both the structural aspects of Leibniz's view of the mind of God (if there be such) as well as some (though not all) of Lewis's views about counterparts, even while preserving Kripkean intuitions that ground and anchor the structure by means of the properties ordinary individuals exemplify at each possible world.

Chapter 14

Natural Numbers

WITH URI NODELMAN[349]

14.1 Philosophical Context

What are the natural numbers? Can the natural numbers number absolutely anything, as Frege assumed, or is it sufficient, for the needs of science, if the natural numbers can only count the objects that might be in the natural world? Is there an infinity of natural numbers and, if so, can this be established in some way other than by stipulation? What primitive notions and axioms do we need to prove the basic postulates governing natural numbers? Must we assume primitive mathematical notions and mathematical axioms for the proof of these postulates, or can we define the notions involved in the postulates non-mathematically and derive the postulates from more general principles?

We try to answer these questions in the present chapter. We begin by recalling Remark (309), in which we distinguished *natural* from *theoretical* mathematics. In that Remark, it was noted that true ordinary statements of number ("there are eight planets", etc.) constitute a body of pretheoretical claims that are assertible without assuming any explicit mathematical theory of numbers. These statements of number are therefore part of *natural* mathematics. They are to be distinguished from the statements mathematicians make when asserting either the axioms or theorems of some mathematical theory, such as Dedekind/Peano number theory, set theory, group theory, etc. These latter are part of *theoretical* mathematics.

In this chapter, we shall analyze the natural numbers as a part of natural mathematics rather than theoretical mathematics. As we shall see, there is a tight connection between the natural numbers and true statements about number assertible in pre-theoretic, ordinary language. To articulate this connection, we shall adapt some of the techniques that Frege developed to define the natural numbers and to derive the postulates and principles that govern them. However, whereas Frege thought that the natural numbers are cardinal numbers that count *all* the objects that exemplify a property, our analysis takes them to have a more limited range of application; in the present system, the natural numbers are natural cardinals that count the *discernible* objects, as these were defined in Section 9.11.3. Our view is that a precondition of counting is that the things being counted are discernible.[350] We would argue that *this is sufficient for the needs of the natural sciences*. On our analysis, the natural numbers emerge as abstractions from the patterns of properties exemplified by discernible objects – this is what makes them *natural* numbers.

Moreover, we shall extend the present theory of abstract objects with an intuitive, logico-metaphysical axiom that uses no mathematical primitives. The resulting theory is capable of (a) *defining* the natural numbers as abstract objects, and (b) *deriving* the number-theoretic postulates as theorems. In particular, the central notions of second-order Peano Arithmetic ('PA2') will be defined and the standard postulates of that theory will be derived as theorems, including a Recursion Theorem that grounds all the (primitive) recursive functions in relation comprehension.

Consequently, this work is to be contrasted with that in Chapter 15, where we use rather different methods to analyze the language, axioms, and theorems of theoretical mathematics. Theoretical mathematics includes the various number theories, set theories, algebras and group theories, etc., and so includes any theory that assumes mathematical primitives and axioms. Thus, PA2 will make an appearance in the next chapter as well as in this one, though in the next chapter, its primitive notions will *not* be analyzed using Frege's methods, but rather using techniques that allow us to analyze any system of theoretical mathematics.

In this chapter, however, we formulate answers to the questions posed at the outset. Moreover, we show that the existence of an infinite cardinal can be proved without appealing to *any* mathematical primitives or mathematical axioms. Thus, we shall be able to answer what Heck (2011, 152) calls the fundamental epistemological question of the philosophy of arithmetic, namely, "What is the basis of our knowledge of the infinity of the series of natural numbers?" Our answer will be that such knowledge can be derived from the logico-metaphysical principles governing abstract objects generally; no mathematics has to be assumed.

[350] As we'll see later, one can 'count', in some sense, indiscernibles by using logically-defined 'numerical' quantifiers, but these won't correspond to statements of number. We'll discuss this further in (847) below.

The views developed in this and the next chapter together bear some similarity to the famous quip attributed to Leopold Kronecker by his students, namely, that "God made the whole numbers; all the others are the work of men".[351] But instead of appealing to God as the source of the natural numbers, we find their origins in the principles governing abstract individuals.[352]

(**740**) **Remark:** Second-Order Peano Arithmetic and the Dedekind/Peano Postulates. As we just mentioned, one of the goals of this chapter is to *derive* PA2. To do this, we'll focus first, and at length, on deriving the Dedekind/Peano postulates following Frege. Both Dedekind 1888 and Peano 1889 contain statements of these basic postulates. We formulate them below, though the presentation doesn't exactly match that of either author.[353] To assert the postulates in a logically perspicuous way, three primitives are needed. They are:

- the individual *Zero*, denoted by the constant 0
- the property *being a number*, denoted by the unary relation constant N
- the binary relation *successor of*, denoted by the binary relation constant S, or the converse relation *predecessor of*, denoted by the binary relation constant P

So in what follows, we understand formulas of the form Nx as asserting that x exemplifies *being a number*, formulas of the form Sxy as asserting that x is an *an immediate successor of y*, and formulas of the form Pxy as asserting that x is

[351] This is reported by H. Weber, who attributes it to Kronecker in Weber 1893 (15). The passage in Weber attributing the quote to Kronecker is:

> Manche von Ihnen werden sich des Ausspruchs erinnern, den er in einem Vortrag bei der Berliner Naturforscher-Versammlung im Jahre 1886 that [sic] "Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk".

As far as we've been able to discover, Kronecker never published this particular statement.

[352] Similarly, we avoid Bishop's (1967, 2) endorsement of Kronecker's appeal to God (as well as to the primitive concepts of a *unit* and *adjunction*):

> The primary concern of mathematics is number, and this means the positive integers. ... The positive integers and their arithmetic are presupposed by the very nature of our intelligence and, we are tempted to believe, by the very nature of intelligence in general. The development of the positive integers from the primitive concept of the unit, the concept of adjoining a unit, and the process of mathematical induction carries complete conviction. In the words of Kronecker, the positive integers were created by God.

[353] Dedekind 1888 (§71) stipulates what must obtain for a set N to be 'simply infinite' or inductive, namely, N must contain an element 1 and be a subset of some set S for which there is a function φ on S such that (a) φ maps N into N, (b) N is the minimal closure of the unit set $\{1\}$ in S under φ, (c) 1 is not the value of φ for any member of N, and (d) φ is a one-to-one function. See Reck 2016 (Section 2.2), where he notes "it is not hard to show that these Dedekindian conditions are a notational variant of Peano's axioms for the natural numbers."

an *immediate predecessor of* y. Also, for present purposes, we may read Sxy as *x succeeds y* and read Pxy as *x precedes y*.

We can state the Dedekind/Peano postulates in terms of the above primitives as follows, where m, n, k are restricted variables ranging over the assumed domain of numbers; for each of the axioms below stated in terms of *succeeds*, we give an alternative statement using *precedes*:

1. Zero is a number.
   $N0$

2. Zero doesn't succeed any number.
   $\neg\exists n S0n$

   No number precedes Zero.
   $\neg\exists n Pn0$

3. If a number k succeeds numbers n and m, then $n = m$.
   $\forall n \forall m \forall k(Skn \,\&\, Skm \rightarrow n = m)$

   If numbers n and m precede a number k, then $n = m$.
   $\forall n \forall m \forall k(Pnk \,\&\, Pmk \rightarrow n = m)$

4. Every number is succeeded by some number.
   $\forall n \exists m Smn$

   Every number precedes some number.
   $\forall n \exists m Pnm$

5. Mathematical Induction: If (a) Zero exemplifies F and (b) Fm implies Fn whenever n succeeds m, then every number exemplifies F.
   $F0 \,\&\, \forall n \forall m(Snm \rightarrow (Fm \rightarrow Fn)) \rightarrow \forall n Fn$

   If (a) Zero exemplifies F and (b) Fn implies Fm whenever n precedes m, then every number exemplifies F.
   $F0 \,\&\, \forall n \forall m(Pnm \rightarrow (Fn \rightarrow Fm)) \rightarrow \forall n Fn$

In addition, Boolos (1995, 293; 1996, 275), and Heck (2011, 288) include the following among the Dedekind/Peano postulates:

6. If x succeeds n, x is a number.
   $\forall n \forall x(Sxn \rightarrow Nx)$

   If n precedes x, x is a number.
   $\forall n \forall x(Pnx \rightarrow Nx)$

7. If numbers m and k succeed a number n, then $m=k$.
   $\forall n \forall m \forall k(Smn \,\&\, Skn \rightarrow m=k)$

   If a number n precedes numbers m and k, then $m = k$.
   $\forall n \forall m \forall k(Pnm \,\&\, Pnk \rightarrow m=k)$

All seven postulates will be derived as theorems of the natural numbers in what follows.

As noted previously, our derivation of these postulates will involve some of the methods that Frege developed. More specifically, we adapt some of the methods used in the proof of *Frege's Theorem*, which is the claim that the Dedekind/Peano postulates for number theory are derivable from a single principle (known as *Hume's Principle*) in second-order logic. Frege's Theorem and Hume's Principle are discussed further below, but before we turn to that discussion, it will be useful to have some definitions and a theorem before us.

Once we've established the Dedekind/Peano postulates as theorems, we'll subsequently explore the theory of numbers that emerges within an object-theoretic background. This will lead us to a full-scale derivation of PA2. This latter theory extends the Dedekind/Peano postulates with the recursive principles of addition and multiplication, and a general comprehension principle for asserting the existence of numerical properties for any condition on the natural numbers:

$$n + 0 = n$$
$$n + m' = (n + m)'$$
$$n \times 0 = 0$$
$$n \times m' = n + (n \times m)$$

$\exists F \forall n(Fn \equiv \varphi)$, where φ is any formula of the language of PA2 in which F doesn't occur free.

We'll derive these principles as theorems of object theory.

(**741**) **Definition:** Correlates One-to-One. Let us say that a binary relation R *correlates* properties F *and* G *one-to-one*, written $R |\!: F \xleftrightarrow{1\text{-}1} G$, just in case (a) R, F, and G exist, (b) each object exemplifying F (hereafter F-object) is R-related to a unique G-object, and (c) each G-object is such that a unique F-object is R-related to it:

$$R |\!: F \xleftrightarrow{1\text{-}1} G \;\equiv_{df}$$
$$R{\downarrow} \,\&\, F{\downarrow} \,\&\, G{\downarrow} \,\&\, \forall x(Fx \rightarrow \exists! y(Gy \,\&\, Rxy)) \,\&\, \forall y(Gy \rightarrow \exists! x(Fx \,\&\, Rxy))$$

Note that the three existence conditions at the beginning of the definiens ensure that the definiendum will be true only when terms instancing the definition all have denotations, so that the definiendum acts like atomic formula.[354]

It is important to remember here that when R correlates F and G one-to-one, the definition: (a) implies nothing about whether R relates F-objects to any $\overline{G}$-objects, (b) implies nothing about whether R relates $\overline{F}$-objects to any G-objects, and (c) implies nothing about whether R relates $\overline{F}$-objects to any $\overline{G}$-objects. For all we know, R might relate an F-object to a single G-object and to several $\overline{G}$-objects. Or R might relate an $\overline{F}$-object to both a $\overline{G}$-object and to a G-object. None of these circumstances are ruled out when R correlates F and G one-to-one. It would serve well to give an example of a relation R that correlates F and G one-to-one even though R relates various $\overline{F}$-objects and $\overline{G}$-objects to one another.[355]

> **Example 1**. Suppose that there are exactly seven individuals $a - g$, and that the relevant facts about them are depicted in Figure14.1: a and b are the only F-objects; c and d are the only G-objects; e exemplifies $\overline{F}$; f and g exemplify $\overline{G}$; and Rac, Rag, Rbd, Rec, and Ref. Then $R|: F \overset{1\text{-}1}{\longleftrightarrow} G$.

(**742**) **Remark:** Digression on the Difference with Frege's Definition in 1884. One of the leading ideas in 1884 is that concepts F and G are equinumerous just in case there is a relation R that correlates F and G one-to-one. But students of Frege may observe that our definition of R correlates F and G one-to-

[354]For those who have skipped the discussions of the theory of definitions, the definition that follows in the text is governed by Convention (17.2). That convention tells us that the following definition gives rise to well-formed instances when any terms substitutable for the free variables are so substituted, and when any alphabetic variant of the definiens replaces the definiens. Without this Convention, we would, strictly speaking, have to formulate the definition using metavariables. To see how this reduces cognitive load, suppose Π and Π' are metavariables ranging over unary relation terms and Π^2 is a metavariable ranging over any binary relation term. And suppose Π^2, Π, or Π' don't contain free occurrences of the individual variables ν and ν'. Then, without our Convention,; the definition in the text would have to be formulated as:

$$\Pi^2 |: \Pi \overset{1\text{-}1}{\longleftrightarrow} \Pi' \equiv_{df}$$
$$\Pi^2\downarrow \,\&\, \Pi\downarrow \,\&\, \Pi'\downarrow \,\&\, \forall\nu(\Pi x \to \exists!\nu'(\Pi'\nu' \,\&\, \Pi^2\nu\nu')) \,\&\, \forall\nu'(\Pi' y \to \exists!\nu(\Pi\nu \,\&\, \Pi^2\nu\nu'))$$

By casting the definition in the text with object-language variables, the definition becomes much easier to read and to understand. Moreover, it yields the same identities and truth conditions whenever it is instanced by denoting terms. See the discussion in Remarks (27) and (28).

Note also that the clause $R\downarrow$ is required for the case where F and G are instanced by necessarily false properties. For if F and G are necessarily false, but R is instanced by a non-denoting relation term, the two main conjuncts of the definiens would be true by failure of the antecedent, and thus the definiendum would be derivable despite the fact that the term instancing R doesn't denote. See the discussion in (36).

[355]For the purposes of the following illustration (and those that we'll construct later in the chapter), we assume that F and G do not overlap—i.e., there is no object x such that $Fx \,\&\, Gx$. Also, we focus our attention on a few particular $\overline{F}$-objects and a few particular $\overline{G}$ objects. That is, we do not concern ourselves with depicting the overlap between $\overline{F}$ and $\overline{G}$. And we do not bother to say explicitly that e is a $\overline{G}$ object and that f and g are $\overline{F}$ objects.

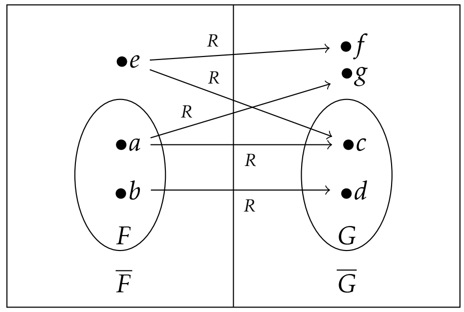


Figure 14.1: R correlates F and G one-to-one.

one in (741) differs from the one Frege gives in the *Grundlagen* of 1884. Our definition is weaker than his. In 1884 (§§71,72), Frege takes R to correlate F and G one-to-one just in case the following four conditions hold:

(.1) each F-object is R-related to some G-object, i.e., $\forall x(Fx \rightarrow \exists y(Gy \,\&\, Rxy))$

(.2) each G-object has an F-object R-related to it, i.e., $\forall y(Gy \rightarrow \exists x(Fx \,\&\, Rxy))$

(.3) R is functional, i.e., $\forall x \forall y \forall z(Rxy \,\&\, Rxz \rightarrow y = z)$

(.4) R is one-to-one, i.e., $\forall x \forall y \forall z(Rxz \,\&\, Ryz \rightarrow x = y)$

This is stronger than the definition of *correlates one-to-one* that we introduced in (741). The definition in (741) doesn't require R to be functional globally and doesn't require R to be one-to-one globally. For example, we saw that the relation R in Figure 14.1 correlates F and G one-to-one in the sense of (741). But R fails to be functional globally, since it relates e to both c and f and a to both c and g. And R fails to be one-to-one globally, since it relates both e and a to c. So, R in Figure 14.1 fails to correlate F and G one-to-one in Frege's sense, since it fails the first two clauses of his definition.

Despite this difference, our weaker definition is all that is needed to derive the Dedekind/Peano postulates — as long as R meets the definition of (741), there is, intuitively, a one-to-one correspondence between the F-objects and the G-objects. If we restate the definition while temporarily assuming some mathematics just for the purpose of illustration, then we could say that our notion of *correlates one-to-one* is sufficient to partition the domain of properties into equivalence classes of equinumerous properties.[356] Finally, Frege's definition of *equinumerosity* changes slightly in 1893 (see Heck 1993, 586; 2011,

[356] Clearly, if R correlates F and G one-to-one in Frege's sense, then R correlates F and G in our sense. Moreover, if R correlates F and G one-to-one in our sense, there is an R' that correlates F and G in Frege's sense. Choose R' to be $[\lambda xy\, Rxy \,\&\, Fx \,\&\, Gy]$, i.e., the restriction of R to those F-objects in its domain and to those G-objects in its range. So, while our definition of what it takes for some relation R to correlate F and G one-to-one is weaker, the same Fs and Gs can always be

47). In what follows, we'll compare our definition with his 1893 definition in footnotes rather than in a Remark.

(**743**) **Definitions:** Relation R Maps F to G. We now say that (.1) R *maps* F *to* G just in case (a) R, F, and G exist, and (b) each F-object is R-related to a unique G-object:

$$(.1)\ \ R|\!:F \longrightarrow G \ \equiv_{df}\ R{\downarrow} \ \&\ F{\downarrow} \ \&\ G{\downarrow} \ \&\ \forall x(Fx \rightarrow \exists!y(Gy \ \&\ Rxy))$$

Cf. Frege 1893, §38.[357]

A caution analogous to the one for definition (741) is in order. Though we'll define functions, domains, and codomains in Section 14.8.2, note that the definition of $R|\!:F \longrightarrow G$ doesn't imply that that R is a function from domain F to codomain G, for it allows R to additionally relate both F-objects to other things (including $\overline{G}$-objects), and relate $\overline{F}$-objects to other things (including G and $\overline{G}$-objects). An example will make it clear why we've posted this caution:

> **Example 2.** Suppose that there are exactly seven individuals a – g, and that the relevant facts about them are depicted in Figure 14.2: a and b are the only F-objects; c and d are the only G-objects; e exemplifies $\overline{F}$; f and g exemplify $\overline{G}$; and Rad, Rag, Rbd, Rec, and Ref. Then $R|\!:F \longrightarrow G$.

So it is important to keep in mind that when $R|\!:F \longrightarrow G$, the definition tells us only some of what R may do and, indeed, only some of what R may do with respect to F-objects and with respect to G-objects.

We next say (.2) R maps F to G *one-to-one*, written $R|\!:F \xrightarrow{1\text{-}1} G$, just in case R maps F to G and for any F-objects x and y and any G-object z, if R relates both x and y to z, then x is identical to y:

$$(.2)\ \ R|\!:F \xrightarrow{1\text{-}1} G \equiv_{df} R|\!:F \longrightarrow G \ \&\ \forall x \forall y \forall z((Fx \ \&\ Fy \ \&\ Gz) \rightarrow (Rxz \ \&\ Ryz \rightarrow x = y))$$

correlated one-to-one by either definition. In other words, both our definition and Frege's induce the same partition on the domain of properties.

[357] To help the reader make the comparison, note that Frege introduced (1893, §37, Definition Γ [2013, 55]) a concept named 'I' under which *single-valued* relations fall; i.e., the concept I takes, as argument, an extension of a relation (i.e., a double value-range) and yields, as value, the True when the argument is the extension of a functional relation, i.e., the relation never maps its arguments to two or more values. So Ip is the True when p is an extension of a functional relation. He then introduced (1893, §38, Definition Δ [2013, 56]) the operator $\rangle$, which intuitively takes, as argument, an extension of a relation (again, a double value-range) and yields, as value, a new relation that relates two extensions just in case the original relation is single-valued and *maps* each member of the first extension to a member of the second extension.

So, intuitively, Frege would say that R maps F to G just in case (a) R is *functional*, as defined in (742.3) above, and (b) R relates each F-object to some G-object, as in (742.1). But, as noted in connection with Frege's 1884 definition, our definition of R maps F to G is weaker since it doesn't require R to be a function globally – R only has to be functional with respect to the objects exemplifying F, i.e., it has to map every such F to a unique G. This will suffice for the same reasons discussed in the previous footnote.

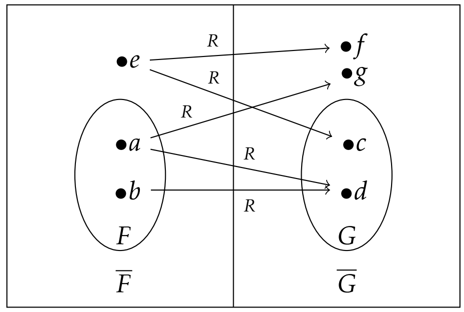


Figure 14.2: R maps F to G.

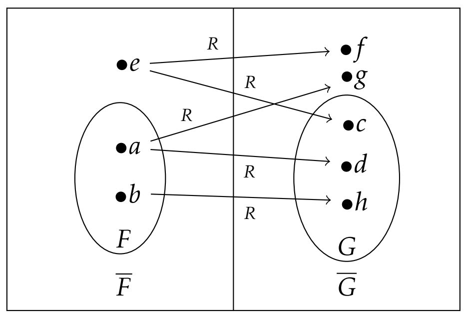


Figure 14.3: R maps F to G one-to-one.

As a check, the reader should verify that R in Example 2 does not map F to G one-to-one. But consider:

> **Example 3**. Suppose that there are exactly eight individuals a – h, and that the relevant facts about them are depicted in Figure 14.3: a and b are the only F-objects; c, d, and h are the only G-objects; e exemplifies $\overline{F}$; f and g exemplify $\overline{G}$; and Rad, Rag, Rbh, Rec, and Ref. Then $R|\!: F \xrightarrow{1\text{-}1} G$.

The reader should verify that in Example 3, R maps F to G one-to-one but does not *correlate* F and G one-to-one.

Next, we say (.3) R maps F *onto* G just in case R maps F to G and every G-object is such that some F-object bears R to it:

(.3) $R|\!: F \underset{\text{onto}}{\longrightarrow} G \equiv_{df} R|\!: F \longrightarrow G \;\&\; \forall y(Gy \rightarrow \exists x(Fx \;\&\; Rxy))$

The reader should verify that R in Example 3 does not map F onto G. But consider:

> **Example 4**. Suppose that there are exactly eight individuals a – h, and that the relevant facts about them are depicted in Figure 14.4:

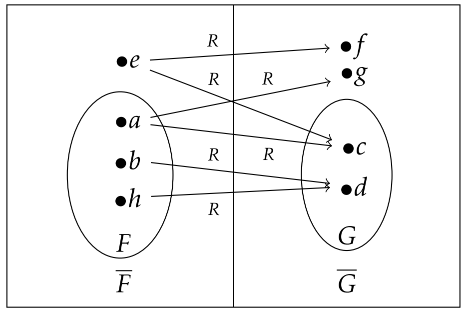


Figure 14.4: R maps F onto G.

> a, b, and h are the only F-objects, c and d are the only G-objects, e exemplifies $\overline{F}$, f and g exemplify $\overline{G}$, and Rac, Rag, Rbd, Rec, Ref, and Rhd. Then $R|: F \underset{\text{onto}}{\longrightarrow} G$.

The reader should also verify that R in Example 4 does not map F to G one-to-one.

Finally, whenever R both maps F to G one-to-one and maps F onto G, we write $R|: F \xrightarrow[\text{onto}]{\text{1-1}} G$:

$$(.4)\quad R|: F \xrightarrow[\text{onto}]{\text{1-1}} G \equiv_{df} R|: F \xrightarrow{\text{1-1}} G \; \& \; R|: F \xrightarrow[\text{onto}]{} G$$

The reader should verify that R in Example 1 maps F onto G one-to-one.

(**744**) **Theorem:** Correlates One-to-One and One-to-One Onto Maps. The following is now derivable from the definitions in (741) and (743), namely, R correlates F and G one-to-one just in case R maps F onto G one-to-one:

$$R|: F \xleftrightarrow{\text{1-1}} G \equiv R|: F \xrightarrow[\text{onto}]{\text{1-1}} G$$

(**745**) **Remark:** Frege's Theorem. Frege's system of 1893, despite being inconsistent, contained one of the most astonishing intellectual achievements in logic and philosophy. This is now known as *Frege's Theorem*, though Frege himself never formulated the result explicitly as a theorem, nor even thought of it as a *result*. Nevertheless, the theorem Frege proved is that the Dedekind/Peano postulates of number theory can be validly derived using only the resources of second-order logic supplemented by a single principle, namely, Hume's Principle.[358]

To state Hume's Principle, two notions are needed. The first is expressed by the definite description *the number of F*s. Frege thought that this description is

[358] See Wright 1983, Heck 1993, and Zalta 2017, for details.

significant because he believed that for every property F, there is a unique individual x that *numbers* F. The other notion needed for Hume's Principle is the *equinumerosity* of properties F and G. Without a definition of equinumerosity, Hume's Principle has an air of triviality to those encountering it for the first time, for it asserts:

> **Hume's Principle**
> The number of Fs is equal to the number of Gs if and only if F and G are equinumerous.

But the air of triviality can be dispelled if equinumerosity is defined in purely logical terms. Let us say that (.1) F and G are *equinumerous*, written $F \approx G$, just in case there exists a relation R that correlates F and G one-to-one:

(.1) $F \approx G \equiv_{df} \exists R(R|\!: F \xleftrightarrow{1\text{-}1} G)$

Cf. Frege 1884 (§§71,72) and 1893 (§40 [2013, 57]).[359] Then, if we abbreviate *the number of Fs* as $\#F$, Hume's Principle may be formally represented as the following non-trivial claim:

(.2) **Hume's Principle**
$\#F = \#G \equiv F \approx G$

From Hume's Principle, (.1), and theorem (744), it follows that $\#F = \#G$ if and only if there is a relation that maps F onto G one-to-one.

We are now in a position to summarize how Frege derived the Dedekind/-Peano postulates in second-order logic from Hume's principle. But before we do, it is worth digressing briefly to say more about how Frege both attempted to define $\#F$ and *derive* Hume's Principle from his theory of extensions. We can simplify the discussion of how Frege thought this could be done by (a) representing Frege's notion *the extension of F* as ϵF and (b) supposing that ϵF is axiomatized by Frege's version of Basic Law V, which for present purposes, may be written as: $\epsilon F = \epsilon G \equiv \forall x(Fx \equiv Gx)$. It is well known that if ϵF is taken as a primitive and systematized by Basic Law V in second-order logic (with *unrestricted* second-order comprehension for properties), the resulting system is subject to Russell's paradox. But Frege was initially unaware of the paradox and so defined $\#F$ as the extension of the property *being an extension of a property equinumerous to F*.

[359] In Remark (742), we discussed the four conditions involved in Frege's 1884 definition of R correlates F and G one-to-one. In 1893 (§40), however, he simplifies the definition as follows: *R correlates F* and *G one-to-one* just in case (a) R maps F to G and (b) the converse of R maps G to F, where *R maps F* to G is defined in §38 and the converse of R is defined in §39. (Recall footnote 357 above, where we mentioned that his definition of R maps F to G requires R to be functional globally.) Though Frege's 1893 definition of *correlates one-to-one* is equivalent to his 1884 definition, Heck shows (1993, 586, note 22; 2011, 47) that it has 'technical advantages' for some of the theorems Frege proves in 1893.

In the present system, we may represent the *being an extension of a property equinumerous to F* as:

$$[\lambda x\, \exists G(x = \epsilon G \,\&\, G \approx F)]$$

But this is not a core λ-expression, as defined in (9.2). The x bound by the λ occurs in encoding position in the matrix $\exists G(x = \epsilon G \& G \approx F)$ — by the Encoding Formula Convention (17.3) and the definition of identity for objects (23.1), x occurs free in (the definiens of) $x = \epsilon G$. So axiom (39.2) doesn't guarantee that the above expression has a denotation. Consequently, the term $\epsilon[\lambda x\, \exists G(x = \epsilon G \,\&\, G \approx F)]$ isn't guaranteed to have a denotation either. But if we put these facts aside temporarily, we could represent Frege's definition of *the number of Fs* as follows:

$$\#F \;=_{df}\; \epsilon[\lambda x\, \exists G(x = \epsilon G \,\&\, G \approx F)]$$

Frege's next step was to derive Hume's Principle, $\#F = \#G \equiv F \approx G$, from Basic Law V. Although this derivation of Hume's Principle loses most of its interest given that Frege's system is inconsistent, the fact is that once Frege had 'derived' Hume's Principle, he then made no further *essential* appeal to Basic Law V when deriving the Dedekind/Peano postulates as theorems; his derivations of these postulates appealed only to Hume's Principle and the theorems of second-order logic (Heck 1993).

It is now known, however, that if one takes $\#F$ instead of ϵF as a primitive and adds Hume's Principle instead of Basic Law V to second-order logic, the resulting system is provably consistent.[360] Frege's Theorem proceeds by constructing the following definitions, in which P stands for *predecessor* and N stands for *being a natural* (or *finite*) *number*:

$$Pxy \;\equiv_{df}\; \exists F \exists z(Fz \,\&\, y = \#F \,\&\, x = \#[\lambda z'\, Fz' \,\&\, z' \neq z])$$

$$P^{*}xy \;\equiv_{df}\; \forall F[\forall z(Pxz \to Fz) \,\&\, \forall x' \forall y'(Px'y' \to (Fx' \to Fy')) \to Fy] \qquad \text{(The ancestral of } P\text{)}$$

$$P^{+}xy \;\equiv_{df}\; P^{*}xy \vee x = y \qquad \text{(The weak ancestral of } P\text{)}$$

$$0 \;=_{df}\; \#[\lambda x\, x \neq x]$$

$$Nx \;\equiv_{df}\; P^{+}0x$$

Thus Frege *defined* the three primitive notions used in the Dedekind/Peano postulates (namely, 0, N, and P), and then *derived* the postulates as theorems of second-order logic supplemented by Hume's Principle.

[360] See Hodes 1984 (138), Burgess 1984 (639), and Hazen 1985 (252). Geach 1976 (446–7) develops the model that the others describe, but doesn't specifically identify it as a model of second-order logic plus Hume's Principle.

Despite this successful derivation, Frege's Theorem, under this reconstruction of taking $\#F$ as primitive and asserting Hume's Principle as an axiom, doesn't accomplish one of Frege's main goals, namely, that of defining the numbers and deriving fundamental number-theoretic principles without appeal to mathematical notions and axioms. That's because Hume's Principle requires a primitive mathematical notion, $\#F$, and the principle thereby becomes a mathematical axiom. By contrast, in what follows, we derive the Dedekind/Peano postulates without mathematical primitives and axioms.

Moreover, there are a number of problems surrounding the methodology of adding principles such as Hume's Principle to second-order logic, some of which were raised by Frege himself. There is now a substantial literature on these problems and it would take us too far afield to delve into the details (see MacBride 2003 for an overview). But it is worth noting one problem that Frege raised, namely, the Julius Caesar problem (1884, §55), which we can state as follows: if the only principle that systematizes $\#F$ is Hume's Principle, i.e., if Hume's Principle is the sole axiom governing *the number of* Fs, then the open formula $\#F = x$ is not defined and the theory of numbers derived from Hume's Principle doesn't provide, for an arbitrary object x, conditions under which $\#F = x$. In particular, the resulting theory offers no conditions for establishing whether the number of planets, say, is identical to Julius Caesar.

Though much has been written about this problem, we shall not spend time working through the literature, since object theory follows a rather different methodology, namely, of presenting separate principles for the existence and identity of theoretical objects.[361] Identity formulas of the form $x = y$ and $F = G$ are defined generally in object theory, and once we identify the #F as an abstract individual, the general object-theoretic principles of identity yield well-defined conditions under which $\#F = x$. The Julius Caesar problem just doesn't arise.

It is important to recognize, however, that Frege's methods can be adapted and applied in object theory by accepting one justifiable limitation, namely, that the natural numbers can only count discernible objects, as these were defined in (273.2). If objects aren't discernible, then we see no reason why the natural numbers should be able to count them. We'll discuss this justifiable limitation explicitly in Remarks (746), (790), (798), and (817). As it turns out, Frege's goals can be achieved by extending object theory with a single axiom that asserts the existence of an ordering relation (*predecessor*) that is definable in non-mathematical but recognizably Fregean terms. Thus, we plan to show that $\#F$, *Precedes*, and 0 can all be defined in object-theoretic terms, and that

[361] This is the methodology used in set theory, for example. In Zermelo-Fraenkel set theory, for example, the existence axioms (Null Set, Pair Set, Unions, Power Set, Infinity, Separation, and Replacement) are distinct from the axiom of the identity of sets (Extensionality).

the Dedekind/Peano postulates and, subsequently, PA2, are derivable in object theory by extending it with the axiom that *Precedes* is a relation.

After this is done, and the foundations of arithmetic are developed in some detail with numerous further definitions and theorems (including one grounding recursively-defined arithmetic functions), we conclude with an interesting result, namely, that, *without any mathematical primitives or axioms*, there is a derivation of the existence of both an infinite cardinal ($\aleph_0$) and an infinite set $\epsilon\mathbb{N}$ (i.e., the extension of the property *natural number*).

14.2 Equinumerosity and Discernible Objects

Throughout the remainder of this chapter, we use u and v as rigid restricted variables ranging over discernible objects. We continue to use x, y, and z as variables ranging over all individuals. Note also that when we compare theorems proved below to theorems found in Frege's work, we assume that our notion of a *property* (i.e., a unary relation) corresponds to Frege's notion of a *concept*. The justification for this comes from Frege himself, who said (1892, 51) that the concepts under which an object falls are its properties. We shall not, however, adopt Frege's analysis of concepts as functions from objects to truth values.

(**746**) **Remark:** Classical Equinumerosity Isn't an Equivalence Condition. One of the keys to Frege's theorem that we haven't discussed is the fact that the equinumerosity of F and G ($F \approx G$), as defined in (745.1), is an equivalence condition in the classical second-order predicate calculus. We leave it as an exercise for the reader to show that in classical second-order logic (without encoding), the condition $F \approx G$ is reflexive, symmetric, and transitive. Frege intuitively relied on this fact, which partitions the domain of properties into equivalence classes of equinumerous properties, to introduce a new object, $\#F$, to represent the class of all properties equinumerous to F.

However, in the present system, classical equinumerosity, as defined in (745.1), provably fails to be an equivalence condition on properties, and this is the first obstacle we must surmount if we are to adapt Frege's methods to object theory. To see why the equinumerosity of F and G fails to be an equivalence condition, recall that it was established in (269) that there are distinct abstract objects that exemplify the same properties:

$$\exists x \exists y(A!x \,\&\, A!y \,\&\, x \neq y \,\&\, \forall F(Fx \equiv Fy)) \qquad (269)$$

From this it follows that $A!$ is not equinumerous to any property, i.e.,

(.1) $\forall G(A! \not\approx G)$

Proof. Given (269), assume a and b are such objects, so that we know $A!a$, $A!b$, $a \neq b$, and $\forall F(Fa \equiv Fb)$. Suppose, for reductio, that $\exists G(A! \approx G)$. Let Q be such a property so that we know $A! \approx Q$. Then, by definition of $\approx$ (745.1), there is a relation, say R, that correlates $A!$ and Q one-to-one, i.e., $R|: A! \xleftrightarrow{1\text{-}1} Q$. So by (744), we know both that R maps $A!$ to Q one-to-one and that R maps $A!$ onto Q. *A fortiori*, R maps $A!$ to Q. The latter fact and our assumption $A!a$ jointly imply that there is a (unique) object, say c, such that both Qc and Rac. So by Rule $\overleftarrow{\beta}$C (184.2.a) and the facts that $[\lambda z\, Rzc]{\downarrow}$ and $a{\downarrow}$, it follows that $[\lambda z\, Rzc]a$. But, since a and b, by hypothesis, exemplify the same properties, $[\lambda z\, Rzc]b$. So by Rule $\overrightarrow{\beta}$C (184.1.a), Rbc. But this contradicts the fact R maps $A!$ to Q one-to-one, for we now have $A!a$, $A!b$, Qc, Rac and Rbc, which by the second conjunct of (743.2) implies $a = b$. Contradiction.

In particular, then, it follows that $A! \not\approx A!$. This result should come as no surprise; since $\approx$ is defined in terms of exemplification patterns of properties and there are abstract objects that are indistinguishable by classical exemplification patterns (269), one should expect that no relation can correlate $A!$ one-to-one with itself.

Given that $A! \not\approx A!$, we have established that equinumerosity is not a reflexive condition:

(.2) $\exists F(F \not\approx F)$

Since equinumerosity is not reflexive, it is not an equivalence condition. Intuitively, then, equinumerosity doesn't partition the domain of properties into mutually exclusive and jointly exhaustive cells of equinumerous properties. Since the existence of such a partition is essential to Frege's method of abstracting out a distinguished object, $\#G$, that numbers all and only the properties equinumerous to G, it will not do us much good to define $\#G$ as $\imath x(A!x \& \forall F(xF \equiv F \approx G))$. Fortunately, however, there is a notion of equinumerosity in the neighborhood which does what Frege's method requires.

This new development will draw upon the theorems about discernible ($D!$) objects (273.2) and the relation $=_D$ of identical discernibles (273.17). The definition of $D!$ ensures that an object x is discernible if x can be identified by the pattern of its exemplifications—i.e., any object y for which $\forall F(Fy \equiv Fx)$ is such that $y = x$. These theorems are presented in (273) and the reader should have a good grasp of the principles governing discernible objects in what follows. Morever, since (273.8) ($D!x \rightarrow \Box D!x$) and (340) imply that $D!x$ is a rigid restriction condition on objects, we shal use u and v as rigid restricted variables ranging over discernible objects to state those theorems. Thus, the definiens of the relation $=_D$ in definition (273.17) becomes expressible as $[\lambda uv\, u = v]$. And we can more simply express theorem (273.34) as $[\lambda x\, x = u]{\downarrow}$, and express theorem (273.35) as $u \neq v \to [\lambda z\, z = u] \neq [\lambda z\, z = v]$. These facts allow to (a) define a notion of equinumerosity that holds whenever there is a relation that correlates, one-to-one, the discernible objects that exemplifying F with those exemplifying G, and then (b) prove that the resulting notion of equinumerosity is an equivalence condition on properties.

(**747**) **Theorem** and **Definitions:** Equinumerosity with respect to Discernible Objects. Using our rigid restricted variables u and v to range over discernible objects, we first derive the special unique existence quantifier for discernible objects. It is provable that (.1) *there is a unique discernible object such that* φ if and only if (a) there is a discernible object such that φ, and (b) every discernible object such that φ is identical to it:

(.1) $\exists! u\varphi \;\equiv\; \exists u(\varphi \,\&\, \forall v(\varphi^{v}_{u} \to v = u))$

Thus, only discernible objects can serve as witnesses for this unique existence quantifier.[362]

Using this special unique existence quantifier, we say that (.2) binary relation R *correlates*$_D$ F and G *one-to-one*, written $R|\!: F \xleftrightarrow{1\text{-}1}_D G$, just in case (a) each discernible F-object is R-related to a unique discernible G-object, and (b) each discernible G-object is such that a unique discernible F-object is R-related to it:

$$\text{(.2)}\quad R|\!: F \xleftrightarrow{1\text{-}1}_D G \;\equiv_{df}$$
$$R\!\downarrow \,\&\, F\!\downarrow \,\&\, G\!\downarrow \,\&\, \forall u(Fu \to \exists! v(Gv \,\&\, Ruv)) \,\&\, \forall v(Gv \to \exists! u(Fu \,\&\, Ruv))$$

Consider:

> **Example 5**. Suppose that there are exactly ten individuals $a - j$, and that the relevant facts about them are depicted in Figure 14.5: $a - e$ and g are the only *discernible* objects; a, b, and e are the only F-objects; c, d, f and g are the only G-objects; h exemplifies $\overline{F}$; i and j exemplify $\overline{G}$; and Rac, Rbd, Ref, Reg, Rhi, and Rhj. Then $R|\!: F \xleftrightarrow{1\text{-}1}_D G$.

The reader should verify that this is an example of the definition.

---

[362] Note that we can abstract out properties from either the definiens or the definiendum. This is due to theorem (273.13), which we can express as $[\lambda u\, \varphi]\!\downarrow$ (for any formula φ), and more generally to theorem (273.13), which we can express as $[\lambda u_1 \ldots u_n\, \varphi]\!\downarrow$ (again, for any formula φ).

To see how, consider the following example of the right-hand condition of (.1). From:

$$\exists v(Gv \,\&\, Rav \,\&\, \forall v'(Gv' \,\&\, Ra' \to v' = v))$$

we can infer:

$$[\lambda u\, \exists v(Gv \,\&\, Ruv \,\&\, \forall v'(Gv' \,\&\, Ruv' \to v' = v))]a$$

by Rule $\overleftarrow{\beta}$C (184.2), since the λ-expression exists by (273.15).

Similar considerations apply to the definiendum. If u occurs free in φ, then $[\lambda u\, \exists! v\varphi]\!\downarrow$, again by (273.13).

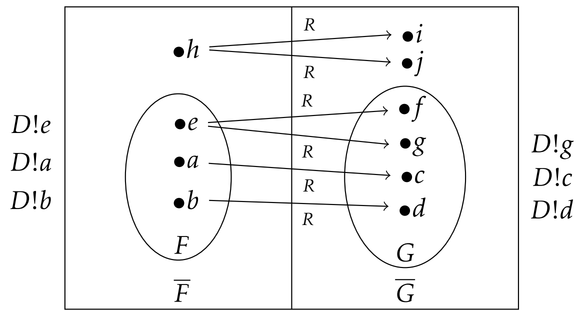


Figure 14.5: R correlates$_D$ F and G one-to-one.

Finally, we say that (.3) properties F and G are *equinumerous with respect to the discernible objects*, or *equinumerous*$_D$, written $F \approx_D G$, just in case some relation correlates$_D$ F and G one-to-one:

(.3) $F \approx_D G \equiv_{df} \exists R(R|: F \xleftrightarrow{1\text{-}1}_D G)$

In the proofs of the theorems that follow, we say that a relation R is a *witness* to the equinumerosity$_D$ of F and G whenever $R|: F \xleftrightarrow{1\text{-}1}_D G$. So R in Figure 14.5 is a witness to the equinumerosity$_D$ of F and G. Also, we sometimes say that F is equinumerous$_D$ *to* G when F and G are equinumerous$_D$.

(**748**) **Theorems:** Equinumerosity$_D$ Partitions the Domain of Properties. It follows that: equinumerosity$_D$ is (.1) reflexive; (.2) symmetric; and (.3) transitive:

(.1) $F \approx_D F$

(.2) $F \approx_D G \rightarrow G \approx_D F$

(.3) $(F \approx_D G \,\&\, G \approx_D H) \rightarrow F \approx_D H$

It also follows that (.4) if F and G are equinumerous$_D$, then a property is equinumerous$_D$ to F iff it is equinumerous$_D$ to G:

(.4) $F \approx_D G \equiv \forall H(H \approx_D F \equiv H \approx_D G)$

(.4) is based on Theorem 25 in Frege 1893 (§61, [2013, 83]).

(**749**) **Definitions:** R Maps$_D$ F to G. We now adapt the notions defined in (743) as follows: (.1) R *maps*$_D$ F *to* G just in case each discernible F-object is R-related to a unique discernible G-object:

(.1) $R|: F \longrightarrow_D G \equiv_{df} R\downarrow \,\&\, F\downarrow \,\&\, G\downarrow \,\&\, \forall u(Fu \rightarrow \exists! v(Gv \,\&\, Ruv))$

Consider:

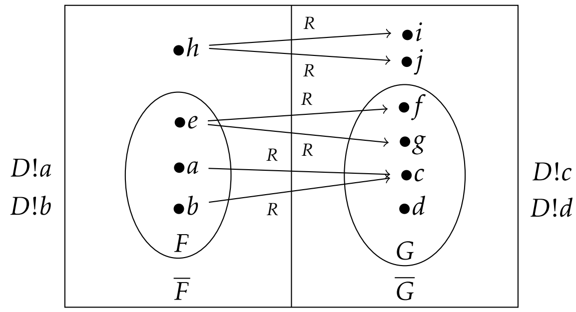


Figure 14.6: R maps$_D$ F to G.

Example 6. Suppose that there are exactly ten individuals $a - j$, and that the relevant facts about them are depicted in Figure 14.6: $a - d$ are the only discernible objects; a, b, and e are the only F-objects; c, d, f and g are the only G-objects; h exemplifies $\overline{F}$; i and j exemplify $\overline{G}$; and Rac, Rbc, Ref, Reg, Rhi, and Rhj. Then $R|\!: F \longrightarrow_D G$.

The following exercises may engender a better understanding:

Exercise 1. Since there are discernible objects (273.5), suppose a is one. Then consider the relation $R = [\lambda xy\, y = a]$, which exists by (273.34). Show that for any F, $R|\!: F \longrightarrow_D D!$.

Exercise 2. Show that the following definition:

$$R|\!: F \longrightarrow_D G \ \equiv_{df}\ R{\downarrow} \,\&\, F{\downarrow} \,\&\, G{\downarrow} \,\&\, R|\!: [\lambda x\, D!x \,\&\, Fx] \longrightarrow [\lambda x\, D!x \,\&\, Gx]$$

is equivalent to (.1). Note that the definiens employs the more general notion: R maps F to G (743.1). Verify that this alternative definition applies in Example 6.

Next, we say (.2) R maps$_D$ F to G *one-to-one* just in case R maps$_D$ F to G and for any discernible F-objects t and u and discernible G-object v, if t bears R to v and u bears R to v, then t is identical to u:

$$(.2)\ \ R|\!: F \xrightarrow{\text{1-1}}_D G \ \equiv_{df}$$
$$R|\!: F \longrightarrow_D G \,\&\, \forall t \forall u \forall v((Ft \,\&\, Fu \,\&\, Gv) \to (Rtv \,\&\, Ruv \to t = u))$$

Consider:

Example 7. Suppose that there are exactly 11 individuals $a - k$, and that the relevant facts about them are depicted in Figure 14.7: $a - d$ and k are the only discernible objects; a, b, and e are the only F-objects; c, d, f, g, and k are the only G-objects; h exemplifies $\overline{F}$; i and j exemplify $\overline{G}$; and Rac, Rbc, Ref, Reg, Rhi, and Rhj. Then $R|\!: F \xrightarrow{\text{1-1}}_D G$.

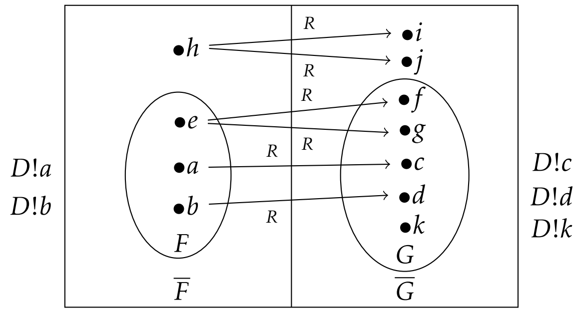


Figure 14.7: R maps$_D$ F to G one-to-one.

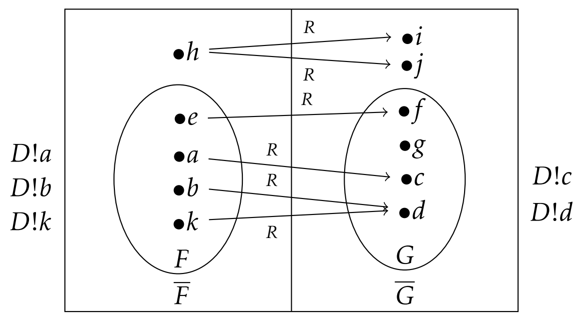


Figure 14.8: R maps$_D$ F onto G.

The following may prove useful:

Exercise 3. Show that the following definition is equivalent to (.2):

$$R|\colon F \xrightarrow{\text{1-1}}_D G \equiv_{df} F{\downarrow} \mathbin{\&} G{\downarrow} \mathbin{\&} R|\colon [\lambda x\, D!x \mathbin{\&} Fx] \xrightarrow{\text{1-1}} [\lambda x\, D!x \mathbin{\&} Gx]$$

Note that this definiens employs the more general notion: R maps F to G one-to-one (743.2). Verify that this alternative definition applies in Example 7.

Next, we say (.3) R maps$_D$ F *onto* G just in case R maps$_D$ F to G and every discernible G-object is such that some discernible F-object bears R to it:

(.3) $R|\colon F \underset{\text{onto}}{\longrightarrow}_D G \equiv_{df} R|\colon F \longrightarrow_D G \mathbin{\&} \forall v(Gv \rightarrow \exists u(Fu \mathbin{\&} Ruv))$

Consider:

Example 8. Suppose that there are exactly 11 individuals a – k, and that the relevant facts about them are depicted in Figure 14.8: a – d and k are the only discernible objects; a, b, e and k are the only F-objects; c, d, f, and g are the only G-objects; h exemplifies $\overline{F}$, i and j exemplify $\overline{G}$; and Rac, Rbd, Rkd, Ref, Rhi, and Rhj. Then $R|\colon F \underset{\text{onto}}{\longrightarrow}_D G$.

The following may prove useful:

> **Exercise 4**. Show that the following definition is equivalent to (.3):
>
> $$R|: F \xrightarrow[\text{onto}]{}_D G \equiv_{df} F{\downarrow} \,\&\, G{\downarrow} \,\&\, R|: [\lambda x\, D!x \,\&\, Fx] \xrightarrow[\text{onto}]{} [\lambda x\, D!x \,\&\, Gx]$$
>
> Note that the definiens employs the more general notion: R maps F onto G (743.3). Verify that this alternative definition applies in Example 8.

Finally, we use the following notation whenever F maps$_D$ F onto G one-to-one:

(.4) $R|: F \xrightarrow[\text{onto}]{\text{1-1}}_D G \equiv_{df} R|: F \xrightarrow{\text{1-1}}_D G \,\&\, R|: F \xrightarrow[\text{onto}]{}_D G$

(**750**) **Theorem:** Correlates$_D$ One-to-One and One-to-One Onto Maps$_D$. It is now easy to show that R correlates$_D$ F and G one-to-one if and only if R maps$_D$ F onto G one-to-one:

$$R|: F \xleftrightarrow{\text{1-1}}_D G \equiv R|: F \xrightarrow[\text{onto}]{\text{1-1}}_D G$$

The reader may find the following useful:

> **Exercise 5**. Using the definitions in the Exercises to (749), show that R correlates$_D$ F and G one-to-one if and only if F and G both exist and R correlates $[\lambda x\, D!x \,\&\, Fx]$ and $[\lambda x\, D!x \,\&\, Gx]$ one-to-one, i.e.,
>
> $$R|: F \xleftrightarrow{\text{1-1}}_D G \equiv F{\downarrow} \,\&\, G{\downarrow} \,\&\, R|: [\lambda x\, D!x \,\&\, Fx] \xleftrightarrow{\text{1-1}} [\lambda x\, D!x \,\&\, Gx].$$

(**751**) **Theorems:** Equinumerosity$_D$ and Empty Properties. The following are modally-strict facts about the equinumerosity$_D$ of empty properties: (.1) if no discernible objects exemplify F and none exemplify H, then F and H are equinumerous with respect to the discernible objects; and (.2) if some discernible object exemplifies F and no discernible object exemplifies H, then F and H aren't equinumerous with respect to the discernible objects:

(.1) $(\neg\exists u Fu \,\&\, \neg\exists v Hv) \rightarrow F \approx_D H$

(.2) $(\exists u Fu \,\&\, \neg\exists v Hv) \rightarrow \neg(F \approx_D H)$

These facts will prove useful later, when we define #G and show that it is not a strictly canonical object.

(**752**) **Theorem and Definition:** Being an F That Is Not Identical to u. Where u is any discernible object, it is straightforward to establish that (.1) the property *being F but not identical to u* exists:

(.1) $[\lambda z\, Fz \,\&\, z \neq u]{\downarrow}$

This holds for any property F and discernible object u. Note that by our conventions in (339.2), the λ-expression in (.1) abbreviates $[\lambda z\, D!y \,\&\, Fz \,\&\, z \neq y]$.

Thus, we may safely introduce the notation F^{-u} to denote *being F but not identical to u*:

(.2) $F^{-u} =_{df} [\lambda z\, Fz \,\&\, z \neq u]$

Although this definition works as intended in what follows, it is important to understand (a) the behavior of the property Π^{-u} when Π is an empty property term, and (b) the behavior of the property $F^{-\kappa}$ when κ is empty or signifies an abstract object, and (c) the behavior of the property $\Pi^{-\kappa}$ when Π is empty *and* κ is empty or signifies an abstract object. **Exercise**. Explain why, in all three of these cases, the resulting expression is significant but denotes an unexemplified property.

(**753**) **Lemma:** An Equinumerosity$_D$ Lemma. If F and G are equinumerous$_D$, u exemplifies F, and v exemplifies G, then F^{-u} and G^{-v} are equinumerous$_D$:

$$F \approx_D G \,\&\, Fu \,\&\, Gv \rightarrow F^{-u} \approx_D G^{-v}$$

Compare this theorem with the intermediate result marked ϑ just prior to Theorem 87 in Frege 1893 (§95, [2013, 125]).[363]

(**754**) **Lemma:** Another Equinumerosity$_D$ Lemma. If F^{-u} and G^{-v} are equinumerous$_D$, u exemplifies F, and v exemplifies G, then F and G are equinumerous$_D$:

$$F^{-u} \approx_D G^{-v} \,\&\, Fu \,\&\, Gv \rightarrow F \approx_D G$$

Compare Frege 1893, Theorem 66 [2013, 112].[364]

(**755**) **Theorem:** Equinumerous$_D$ Is a Contingent Condition on Some Properties. Intuitively, the fact that F and G are equinumerous$_D$ in one possible world doesn't entail that they are equinumerous$_D$ in another. For example, even if *being a pencil on my desk* and *being a pen on my desk* are in fact equinumerous$_D$, things might have been different. But such examples require us to apply the theory by introducing some new, primitive properties. So the question is, can

[363] There are several differences. First, on the basis of Hume's Principle, Frege uses numeral identities instead of equinumerosity claims in the antecedent and consequent, and if we substitute these into the present theorem, we obtain $\#F = \#G \,\&\, Fu \,\&\, Gv \rightarrow \#F^{-u} = \#G^{-v}$. Second, Frege proves the contrapositive, switches the order of the antecedents, and puts everything into conditional form. Thus, he proves: $Fu \rightarrow (Gv \rightarrow (\#F^{-u} \neq \#G^{-v} \rightarrow \#F \neq \#G))$. Note also that Frege uses c and b, respectively, where we use u and v, and he uses v and u, respectively, where we use $\#F$ and $\#G$.

[364] The present theorem differs from Frege's Theorem 66 only by two applications of Hume's Principle: in Frege's Theorem, $\#F^{-u} = \#G^{-v}$ is substituted for $F^{-u} \approx_D G^{-v}$ in the antecedent, and $\#F = \#G$ is substituted for $F \approx_D G$ in the consequent.

we prove that there are properties that are possibly equinumerous$_D$ and possibly not, without applying the theory?

In object theory, we can establish this by way of a preliminary proof that (.1) there are properties F and G such that it is possibly the case that both (a) F and G are equinumerous$_D$ but (b) possibly not:

(.1) $\exists F \exists G \Diamond(F \approx_D G \,\&\, \Diamond\neg F \approx_D G)$

Clearly, by (165.11) and a Rule of Substitution, this implies $\exists F \exists G(\Diamond F \approx_D G \,\&\, \Diamond\neg F \approx_D G)$. So, for some values of F and G, the condition $F \approx_D G$ is provably a contingent condition.

It is also important to observe a variation on this contingency. Namely, (.2) there are properties F and G such that it is possibly the case that both (a) *actually exemplifying* F is equinumerous$_D$ to G, but (b) possibly not:

(.2) $\exists F \exists G \Diamond([\lambda z\, \mathcal{A}Fz] \approx_D G \,\&\, \Diamond\neg[\lambda z\, \mathcal{A}Fz] \approx_D G)$

This analogously implies $\exists F \exists G(\Diamond[\lambda z\, \mathcal{A}Fz] \approx_D G \,\&\, \Diamond\neg[\lambda z\, \mathcal{A}Fz] \approx_D G)$. The distinction between (.1) and (.2) plays an important role in what follows.

(**756**) **Definition:** Material Equivalence with Respect to Discernible Objects. We say that properties F and G are *materially equivalent with respect to discernible objects*, written $F \equiv_D G$, if and only if F and G both exist and are exemplified by the same discernible objects:

$$F \equiv_D G \;\equiv_{df}\; F\downarrow \,\&\, G\downarrow \,\&\, \forall u(Fu \equiv Gu)$$

(**757**) **Lemmas:** Equinumerous$_D$ and Equivalent$_D$ Properties. The following consequences concerning equinumerous$_D$, and materially equivalent$_D$, properties are easily provable: (.1) if F and G are materially equivalent$_D$, then they are equinumerous$_D$; and (.2) if F is equinumerous$_D$ to G and G is materially equivalent$_D$ to H, then F is equinumerous$_D$ to H; and (.3) if F and G are equinumerous$_D$, then for any property H, *actually exemplifying* H is equinumerous$_D$ to F iff it is equinumerous$_D$ to G:

(.1) $F \equiv_D G \to F \approx_D G$

(.2) $(F \approx_D G \,\&\, G \equiv_D H) \to F \approx_D H$

(**758**) ⋆**Theorems:** Equivalence$_D$, Equinumerosity$_D$, and *Actually Being F*. The following theorems prove to be useful. (.1) *actually being* F is materially equivalent to F w.r.t. discernible objects, and (.2) *actually being* F is equinumerous to F w.r.t. discernible objects:

(.1) $[\lambda z\, \mathcal{A}Fz] \equiv_D F$

(.2) $[\lambda z\, \mathcal{A}Fz] \approx_D F$

These theorems are not modally strict.

(**759**) **Theorems:** Facts About Actually Exemplifying F. (.1) It is actually the case that F is equinumerous$_D$ to *actually exemplifying* F and (.2) *actually exemplifying* F is a rigid property:

(.1) $\mathcal{A}(F \approx_D [\lambda z\, \mathcal{A}Fz])$

(.2) $Rigid([\lambda z\, \mathcal{A}Fz])$

The definition of $Rigid(F)$ is in (571).

(**760**) **Theorems:** Conditions for Modal Collapse of Equinumerosity$_D$. (.1) If F is a rigid property (i.e., if it is necessary that every F-thing is necessarily F), then F is equinumerous$_D$ to *actually exemplifying* F:

(.1) $Rigid(F) \rightarrow F \approx_D [\lambda z\, \mathcal{A}Fz]$

With (.1), we can more easily prove, as a modally strict theorem, that (.2) F and G are equinumerous$_D$ if and only if for any property F, the actualization of H is equinumerous$_D$ to F iff it is equinumerous$_D$ to G:

(.2) $F \approx_D G \equiv \forall H([\lambda z\, \mathcal{A}Hz] \approx_D F \equiv [\lambda z\, \mathcal{A}Hz] \approx_D G)$

While it might seem that the above is a trivial consequence of (748.4), there is an important subtlety. To establish the right-to-left direction, one must show that there is no danger of the right-hand side being satisfied trivially, i.e., that no actualization is equinumerous$_D$ to F or G. But, as it turns out, for any property F, there is an actualization that is materially equivalent and thus equinumerous$_D$ to F. We can see this by recalling that the system guarantees the existence of rigidifications of properties (573.3). The actualization of the rigidification of F is equivalent to F and so it must be equinumerous$_D$ to F.

Finally, (.3) if F and G are rigid properties, then the equinumerosity$_D$ of F and G is a modally collapsed condition:

(.3) $(Rigid(F) \,\&\, Rigid(G)) \rightarrow \Box(F \approx_D G \rightarrow \Box F \approx_D G)$

By the definition of rigidity, (.3) implies that if (a) necessarily every F-exemplifier necessarily exemplifies F and (b) necessarily every G-exemplifier necessarily exemplifies G, then necessarily, if F and G are equinumerous$_D$, they are necessarily equinumerous$_D$. (.3) plays an important role when we consider properties like *actually exemplifying* F, e.g., $[\lambda z\, \mathcal{A}Fz]$, since these are rigid, by (759.2).

14.3 Numbering Properties and Natural Cardinals

In this section, we work our way towards a definition of x *numbers* G and then define x *is a natural cardinal*. When we reconstruct Frege's conception of these notions within object theory, we introduce some emendations. We'll see that Frege's conception doesn't anticipate the fact that equivalence classes of equinumerous properties (in Frege's case) and equivalence classes of equinumerous$_D$ properties (in the present theory) vary from world to world. This would give rise to different numbers at different worlds. If there were, say, 2 discernible G things at one possible world and 2 discernible G things at a different possible world, we should be able to say that the same number numbers G at both worlds. But since G will be a member of distinct equivalence classes of equinumerous$_D$ properties at distinct worlds, the object that numbers G at the former world will be distinct from the object that numbers G at the latter world. This runs contrary to our general understanding of how natural cardinals work and it shows that Frege's picture doesn't generalize in a modal context (cf. Panza 2018, 99–100, for a related point).[365]

With a minor adjustment to Frege's conception, however, we can ensure not only that there is one group of abstract objects that both number properties and serve as natural cardinals, but also that these same objects can correctly number properties across modal contexts. We've tried, in the first instance, to develop number theory using only primitive modalities, without invoking the notion of *possible world* defined in Chapter 12. However, it is occasionally helpful to develop intuitive remarks in which we appeal to the notion of a possible world (e.g., in Remark (761) below). This is for exegetical purposes only, to prepare the ground and to help one better understand the issues involved.[366]

(**761**) **Remark:** How to Adapt Frege's Conception So That it Works in a Modal Setting. In 1884 (§68), Frege defines:

- *the Number which belongs to* the concept F is the extension of the concept *equinumerous to the concept F*

In Zalta 1999 (630), this definition was captured in object theory by saying x *numbers* G if and only if x is an abstract object that encodes all and only

[365]Panza (2018) takes as a premise the fact that, for any property G, it is contingent as to which object is $\#G$ given that it is contingent as to which first-order properties exemplify the second-order property: *being a first-order property equinumerous to F*. His conclusion is that $\#G$ isn't therefore a logical object. There are two reasons why we don't draw the same conclusion. (a) To obtain the number of Gs, we abstract over the properties F whose (rigid) *actualizations* are equinumerous to G (w.r.t. discernible objects). This eliminates one of the contingencies that Panza is concerned about. Moreover, (b) we see no reason why *logico-metaphysical* objects can't be abstracted from contingent patterns in the natural world.

[366]See Cook 2016 for an entirely different method that develops Frege's theory of numbers in a modal setting.

the properties F that are equinumerous to G with respect to ordinary objects ($\approx_E$).[367] And, in the current reconstruction of Frege's view, with our focus being on numbering discernible objects instead of just the ordinary objects, we shall replace equinumerosity w.r.t. ordinary objects ($\approx_E$) by equinumerosity w.r.t. discernible objects ($\approx_D$) and formally stipulate:

(ϑ) $Numbers(x,G) \equiv_{df} A!x \,\&\, \forall F(xF \equiv F \approx_D G)$

Though the definition in Zalta 1999 and (ϑ) are analogous to Frege's, both of these definitions give rise to a problem in a modal setting: they imply that different cardinal numbers emerge in different modal contexts. We can restate this problem by speaking intuitively in terms of possible worlds: the above definitions imply that the cardinals at one possible world will be distinct from the cardinals at other possible worlds.

We'll frame the problem in terms of discernibles. The issue isn't that some object, say x, numbers G at possible world w_1 and some different object, say y, numbers G at possible world w_2. That is only to be expected, since G might be exemplified by two discernible objects in w_1 and three discernible objects in w_2. Rather, the problem is that G might be exemplified by two discernible objects in both w_1 and w_2, but the object that numbers G in w_1 is not identical to the object that numbers G in w_2. To see why, note first that (ϑ) gives rise to the following necessary equivalence:

$$\Box(Numbers(x,G) \equiv (A!x \,\&\, \forall F(xF \equiv F \approx_D G))$$

Intuitively, then, the central material biconditional should hold at every possible world. But consider the scenario depicted in Figure 1, which shows only the discernible objects that exemplify P, Q, and R in possible worlds w_1 and w_2. Clearly, this scenario illustrates (755.1), since P and Q are equinumerous$_D$ in w_1 but not in w_2. Now, intuitively, the number of Ps at w_1 should be identical to the number of Ps at w_2, since there are exactly two objects exemplifying P at both w_1 and w_2. But suppose x numbers P at w_1. Then, by (ϑ), x encodes all and only the properties equinumerous$_D$ to P at w_1. Hence, x encodes Q as well, though not R. Now suppose y numbers P at w_2. Intuitively, it should be the case that $y = x$. But then, y encodes R, not Q, since R is equinumerous$_D$ to P at w_2 and Q isn't. Since x and y are thus abstract objects that encode different properties, it follows that $y \neq x$. In general, the numbers carved out at w_1 by (ϑ) are different from the numbers carved out at w_2.

[367] Cf. Boolos (1987, 3; 1998, 184), who suggests (a) that Frege implicitly uses a second 'instantiation relation', namely F is *in* x (which Boolos represents as $F\eta x$) and (b) that Frege conceives of the number x that belongs to a concept G as satisfying the condition: for every concept F, $F\eta x$ if and only if F is equinumerous to G. In the present analysis, we use the more general notion of encoding, xF, instead Boolos' notion $F\eta x$, since the latter was introduced only to assert that properties are *in* objects when some equinumerosity condition obtained.

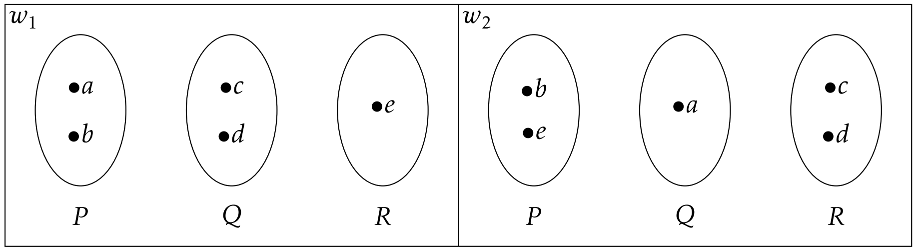


Figure 14.9: Emergence of world-bound natural cardinals.

This is not just a consequence of identifying an object that numbers G as an abstract object that encodes the properties equinumerous$_D$ to G. The problem arises from adopting Frege's conception of numbers in a modal setting. The second-level concept *being equinumerous to P* has different properties falling under it at w_1 and at w_2. So the object that is the extension of this second-level concept at w_1 is different from the object that is the extension of the second-level concept at w_2. Thus, the number that belongs to P at w_1 would be different from the number that belongs to P at w_2, i.e., the number 2 abstracted from w_1 would be distinct from the number 2 abstracted from w_2. Since equivalence classes of equinumerous properties may vary from world to world, the Fregean abstractions on the basis of the simple equivalence condition of equinumerosity (or equinumerous$_D$) will yield *world-bound* numbers.

Now the problem just described didn't have serious consequences for the work in Zalta 1999. The issue was finessed in two ways: (a) $\#G$ was defined in terms of a rigid definite description, i.e., $\imath x Numbers(x, G)$, and (b) the work didn't take pains to distinguish which theorems were derivable without any appeal to contingencies. But the fact remains: Zalta 1999 used (ϑ), and though it was stated in terms of $\approx_E$ instead of $\approx_D$, it too engendered (though it didn't reference) different natural cardinals at different possible worlds.

But in the present attempt to refine Zalta 1999, we want to avoid generating world-bound natural cardinals – our goal is to define a single group of natural cardinals that can be used to count properties at arbitrary possible worlds. We claim that the simplest way to extend Frege's definition to a modal setting unproblematically is to regard the number of Gs, at any world w, as an abstraction over all the properties F for which the discernible objects *actually exemplifying* F are in one-to-one correspondence with the discernible objects exemplifying G at w. This would always use the extensions of properties at the actual world as the reference basis for abstracting the numbers.

Thus, instead of (ϑ), we shall say that x *numbers* G if and only if x is an abstract object that encodes all and only the properties F such that *actually exemplifying F* is equinumerous to G with respect to discernible objects:

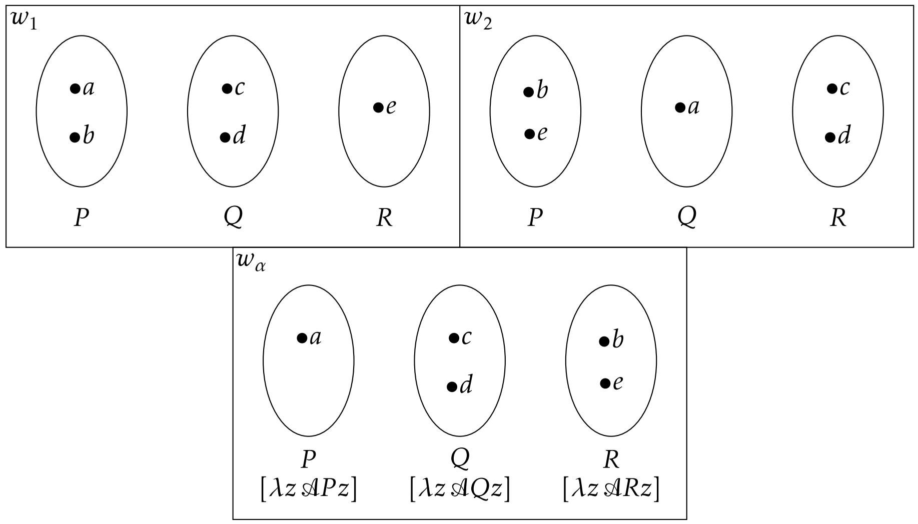


Figure 14.10: Emergence of universal natural cardinals.

(ξ) $Numbers(x, G) \equiv_{df} A!x \,\&\, \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)$

Note that if one were to ignore axiom (45.4), which guarantees the existence of at least two possible worlds, and consider a degenerate modal setting in which the actual world is the only possible world, then (ξ) reduces to (ϑ).

In the usual manner, (ξ) yields a necessary equivalence:

$$\Box(Numbers(x, G) \equiv A!x \,\&\, \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G))$$

But (ξ) and the above equivalence yield *universal* cardinals. To see this more clearly, we add the actual world to our picture, as in Figure 14.10. Now, intuitively, as in the previous figure, any object that numbers P at w_1 should be identical to the object that numbers P at w_2, since there are exactly two discernible objects exemplifying P at both w_1 and w_2. We verify this over the next two paragraphs by referencing (ξ).

Suppose x numbers P at w_1. Then, by (ξ), x encodes all and only the properties F such that $[\lambda z\, \mathcal{A}Fz]$ is equinumerous$_D$ to P at w_1. Inspection shows that x therefore encodes both Q and R, since $[\lambda z\, \mathcal{A}Qz]$ and $[\lambda z\, \mathcal{A}Rz]$ are both equinumerous$_D$ to P at w_1, i.e., there is a one-to-one correspondence from the discernible objects exemplifying $[\lambda z\, \mathcal{A}Qz]$ to the discernible objects exemplifying P at w_1, and similarly from the discernible objects exemplifying $[\lambda z\, \mathcal{A}Rz]$.

Now suppose y numbers P at w_2. Then, by (ξ), y encodes all and only those Fs such that $[\lambda z\, \mathcal{A}Fz]$ is equinumerous$_D$ to P at w_2. But inspection shows that y therefore encodes Q and R, since these are, again, the only such Fs for which *actually being* F is equinumerous$_D$ to P at w_2. If we assume that these are the only encoding facts available, then $x = y$; they are abstract objects that encode the same properties. Not only is the object that numbers P at w_1 identical to the object that numbers P at w_2, but the object that numbers P at w_1 is identical to the object that numbers R at w_2. And so on. In this manner, there will emerge from (ξ) a single group of *universal* natural cardinals and that is why we shall stipulate (ξ) in what follows.

One final point. In 1884, Frege also defines:

> n is a Number if and only if there exists a concept such that n is the Number which belongs to it (§72).

Frege's definition becomes ambiguous in the present context. Once we've defined:

$$\#G =_{df} \iota x Numbers(x, G)$$

one can define a natural cardinal in one of two ways:

$$NaturalCardinal(x) \equiv_{df} \exists G(x = \#G)$$

$$NaturalCardinal(x) \equiv_{df} \exists G(Numbers(x, G))$$

We'll show that the two definientia $\exists G(x = \#G)$ and $\exists G(Numbers(x, G))$ are in fact *strictly* equivalent, though it is important to recognize that the *open* formulas $x = \#G$ and $Numbers(x, G)$ are not! The biconditional $x = \#G \equiv Numbers(x, G)$ is *not* a modally strict theorem of object theory; it *is* a theorem but it rests on a contingency. To see this, fix G and fix the world of evaluation to be w. Then no matter how many discernible objects fall under G at w, the abstract object, say a, that satisfies the left condition ($x = \#G$) will invariably be the abstract object that in fact numbers G at the actual world. But the object, say b, that satisfies the right condition ($Numbers(x, G)$) depends on how many discernible objects fall under G at w and so b may differ from a. Thus the two conditions are guaranteed to be equivalent only at the actual world; that's why $x = \#G \equiv Numbers(x, G)$ is a theorem but not a modally strict theorem.

By contrast, the biconditional $\exists G(x = \#G) \equiv \exists G(Numbers(x, G))$ is a modally strict theorem. We'll return to this fact on occasion as we develop the theory of natural cardinals and natural numbers.

(**762**) **Definitions:** Numbering a Property. Consequently, we officially define: x *numbers* (the discernible objects exemplifying) G if and only if x is an abstract object, G exists, and x encodes just the properties F such that *actually exemplifying* F is equinumerous to G w.r.t. discernible objects. Formally:

$$Numbers(x, G) \ \equiv_{df} \ A!x \,\&\, G\downarrow \,\&\, \forall F(xF \equiv [\lambda z \, \mathcal{A}Fz] \approx_D G)$$

As noted in the previous Remark (761), the use of the actuality operator in the definition of $Numbers(x, G)$ constitutes a small but important modification to Frege's definition. We can render the definition intuitively as: for any world w, the abstract object that numbers G at w encodes those properties F such that the discernible objects exemplifying F at the actual world correspond one-to-one with the discernible objects exemplifying G at w. By our theory of, and conventions for, definitions, and the Rule ≡S of Biconditional Simplification, it follows, for any property term Π, that $\Pi\downarrow \rightarrow (Numbers(x, \Pi) \equiv (A!x \,\&\, \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D \Pi)))$.

(**763**) **Theorems:** Existence of (Unique) Objects that Number Properties. It follows immediately from definition (762) and (strengthened) comprehension for abstract objects that (.1) there exists an object that numbers G; and (.2) there exists a unique object that numbers G:

(.1) $\exists x Numbers(x, G)$

(.2) $\exists! x Numbers(x, G)$

Note here that by RN, (.1) and (.2) are necessarily true. Intuitively, in terms of possible worlds, these necessary truths tell us that at every possible world, there is a (unique) x that numbers G there, i.e., that encodes all and only the properties F such that *actually being* F is equinumerous$_D$ to G there.

(**764**) **Theorem:** Equinumerosity$_D$ and Numbering. (.1) if G and H are equinumerous$_D$, then x numbers G if and only if x numbers H, and (.2) if x numbers both G and H, then G and H are equinumerous$_D$:

(.1) $G \approx_D H \rightarrow (Numbers(x, G) \equiv Numbers(x, H))$

(.2) $(Numbers(x, G) \,\&\, Numbers(x, H)) \rightarrow G \approx_D H$

(**765**) **Theorem:** The Principles Underlying Hume's Principle. It is a modally strict consequence of the foregoing that (.1) if x numbers G and y numbers H, then x is identical to y if and only if G and H are equinumerous with respect to the discernible objects; and (.2) there exists a number that numbers both F and G is and only if F is equinumerous$_D$ to G:

(.1) $(Numbers(x, G) \,\&\, Numbers(y, H)) \rightarrow (x = y \equiv G \approx_D H)$

(.2) $\exists x(Numbers(x, F) \,\&\, Numbers(x, G)) \equiv F \approx_D G$

(.2) directly asserts that the equinumerosity$_D$ of F and G is necessary and sufficient for the existence of an object that numbers both properties, and vice versa. By RN, this modally strict equivalence holds necessarily.

However, Hume's Principle actually says a bit more. If we try to express what Hume's Principle asserts without using terms like #F (which are defined by definite descriptions), then a modally strict version of the principle presents itself, namely (.3) there are objects x and y such that (a) x uniquely numbers F, (b) y uniquely numbers G and (c) x is identical to y, if and only if, F is equinumerous$_D$ to G:

(.3) $\exists x \exists y(Numbers(x,F) \,\&\, \forall z(Numbers(z,F) \rightarrow z = x) \,\&$
$\quad Numbers(y,G) \,\&\, \forall z(Numbers(z,G) \rightarrow z = y) \,\&\, x = y) \equiv F \approx_D G$

(**766**) **Theorem:** Material Equivalence$_D$ and Numbering. It should be clear that if G is materially equivalent to H with respect to the discernible objects, then x numbers G if and only if x numbers H:

$$G \equiv_D H \rightarrow (Numbers(x,G) \equiv Numbers(x,H))$$

(**767**) **Remark:** Numbered Properties and Material Equivalence$_D$. Intuitively, the fact that x numbers both G and H doesn't imply that G and H are materially equivalent with respect to the discernible objects. So the following claim should be *demonstrably false*:

$$(Numbers(x,G) \,\&\, Numbers(x,H)) \rightarrow G \equiv_D H$$

Of course, examples are easy to come by: if one and the same abstract object numbers *being a planet* and *being a human-manufactured artifact on my desk*, it doesn't follow that all and only the discernible objects that exemplify *being a planet* exemplify *being a human-manufactured artifact on my desk*.

But the question is, can one *prove* the negation of the claim displayed above *within* our system without additional assumptions, i.e., can we prove, for some object x and properties G and H, that both $Numbers(x,G)$ and $Numbers(x,H)$ but G is not equivalent$_D$ to H:

$$\exists x \exists G \exists H(Numbers(x,G) \,\&\, Numbers(x,H) \,\&\, \neg G \equiv_D H)$$

As it turns out, the proof of this claim holds under the condition that there are at least two discernible objects, as the next theorem establishes. However, we won't be in a position to show that there are at least two discernible objects until later.

(**768**) **Theorem:** Conditional Fact. If there are at least two discernible objects, then there is an abstract object x and properties G and H such that x numbers both G and H but G and H are not materially equivalent$_D$:

$$\exists u \exists v(u \neq v) \rightarrow \exists x \exists G \exists H(Numbers(x,G) \,\&\, Numbers(x,H) \,\&\, \neg G \equiv_D H)$$

To anticipate, the antecedent of this claim will be derivable. We already know that there is at least one discernible object (227.1), since we know that there is at least one ordinary object and that ordinary objects are discernible. However, to establish that there is a second, distinct discernible object, we will define Zero as an abstract object (782.1), prove that it exists (782.2), show that it is a natural cardinal (783), and then show that natural cardinals are discernible (803.3). Then we'll have established that there are at least two distinct discernibles, since the first discernible is ordinary and the second (Zero) is abstract and so have to be distinct (222.2). The above theorem then holds in virtue of their haecceities, which exist given they are discernibles (273.34); it should be easy to see that there is a single number that numbers both haecceities, even though their haecceities are not materially equivalent$_D$.

(**769**) **Theorem:** Numbering a Property and Modal Collapse. Intuitively, there are properties G and possible worlds w_1 and w_2 such that G is exemplified by n discernible objects at w_1 but isn't exemplified by n discernible objects at world w_2. When that happens, the object that numbers G at w_1 is distinct from the object that numbers G at w_2. Formally, if we use reasoning analogous to that used to establish that equinumerosity$_D$ is a contingent condition (755.2), it is relatively straightforward to show that (.1) there is an object x and a property G such that x numbers G but doesn't necessarily number G:

(.1) $\exists x \exists G(Numbers(x,G) \,\&\, \neg\Box Numbers(x,G))$

This theorem takes on significance when we consider it in conjunction with the necessitation of (763.1). For taken together, they assure us that every property in any modal context is numbered by some object even though different objects may number the property in other modal contexts.

However, it follows generally that (.2) if G is any rigid property, then necessarily, anything that numbers G necessarily numbers G:

(.2) $Rigid(G) \rightarrow \Box\forall x(Numbers(x,G) \rightarrow \Box Numbers(x,G))$

Thus, since the property *actually exemplifying* G, i.e., $[\lambda z \mathcal{A}Gz]$, is rigid (759.2), it follows that (.3) necessarily, anything that numbers $[\lambda z \mathcal{A}Gz]$ necessarily numbers $[\lambda z \mathcal{A}Gz]$:

(.3) $\Box\forall x(Numbers(x,[\lambda z \mathcal{A}Gz]) \rightarrow \Box Numbers(x,[\lambda z \mathcal{A}Gz]))$

Finally, it follows that (.4) it is actually the case that x numbers G if and only if x numbers *actually exemplifying* G:

(.4) $\mathcal{A}Numbers(x,G) \equiv Numbers(x,[\lambda z \mathcal{A}Gz])$

(**770**) **Theorem:** The Number of Gs Exists. We previously established that necessarily, there exists a unique abstract object that numbers G (763.2). So *the number of* Gs exists:

$\imath x Numbers(x, G)\downarrow$

Of course, the definite description rigidly denotes the object that in fact numbers G.

(**771**) **Definition** and **Theorem:** Notation for, and Existence of, the Number of (Discernible) Gs. We now introduce the notation $\#G$ to rigidly refer to *the number of* (discernible) Gs:

(.1) $\#G =_{df} \imath x Numbers(x, G)$

Note that since the existence of G is required by the condition $Numbers(x, G)$, a term of the form $\#\Pi$ will be empty if Π is empty. Clearly, it now follows that the number of Gs exists:

(.2) $\#G\downarrow$

(**772**) **Theorem:** The Number of Gs is Canonical. Here are two theorems that show the number of Gs is (identical to) a canonical object:

(.1) $\#G = \imath x(A!x \,\&\, \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G))$

(.2) $\#G = \imath x(A!x \,\&\, \forall F(xF \equiv F \approx_D G))$

(**773**) **Remark:** The Number of Gs is Not Strictly Canonical. Recall that in (260.2), we stipulated that $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$ is a *strictly* canonical object just in case φ is a rigid condition on properties, i.e., by (260.1), if and only if $\vdash_\Box \forall F(\varphi \to \Box\varphi)$. However, if we let φ be the formula $[\lambda z\, \mathcal{A}Fz] \approx_D G$, then theorem (755.2) has the form $\exists F \exists G \Diamond(\varphi \,\&\, \Diamond\neg\varphi)$. By reasoning analogous to that in Remarks (298) and (326), it follows that φ is not a rigid condition on properties, on pain of system inconsistency. So though $\#G$ is (identical to) a canonical object, it is not (identical to) a strictly canonical object.

Since $\#G$ isn't strictly canonical, some well-known Fregean theorems about this object won't be modally strict. In particular, we'll see that the most obvious way of formulating Hume's Principle won't be modally strict. There is, however, a way of formulating a modally strict versions of Hume's Principle. See the theorems in (775) below.

(**774**) **Lemmas:** Facts About The Number of Gs. We now establish that: (.1) x numbers *actually exemplifying* G iff x is the number of Gs; (.2) the number of Gs numbers *actually exemplifying* G; (.3) the number of Gs is abstract and, for any property F, encodes F iff *being actually* F is equinumerous$_D$ to *actually exemplifying* G; (.4) the number of Gs encodes G; and (.5) if G is rigid, then the number of Gs numbers G:

(.1) $Numbers(x, [\lambda z\, \mathcal{A}Gz]) \equiv x = \#G$

(.2) $Numbers(\#G, [\lambda y\, \mathcal{A}Gy])$

(.3) $A!\#G \,\&\, \forall F(\#GF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D [\lambda z\, \mathcal{A}Gz])$

(.4) $\#GG$

(.5) $Rigid(G) \rightarrow (Numbers(\#G, G))$

In (.3) and (.4), we put the symbol G immediately following the octothorpe in a slightly smaller font size solely for ease of readability. Thus, $A!\#G$ asserts that $\#G$ exemplifies being abstract, while $\#GF$ asserts that $\#G$ encodes F.

Note here that we've done something different in proving the above theorems about $\#G$, namely, that we've proved *modally strict* theorems having the above forms even though $\#G$ is not strictly canonical. In previous chapters, the analogous theorems (i.e., having the above form) governing objects that aren't strictly canonical were not modally strict. Consequently, it is important to observe the subtle differences in the above theorems which yield the modal strictness, such as the places where $[\lambda z\, \mathcal{A}Gz]$ occurs instead of G. Our goal, remember, is to build natural cardinals and natural numbers that are completely general and that can number properties no matter what the modal context. See again Remark (761).

(**775**) **Theorem:** Modally Strict Version of Hume's Principle. Whenever F and G are rigid properties, then the number of Fs is identical to the number of Gs if and only if F is equinumerous$_D$ to G:

$$(Rigid(F) \,\&\, Rigid(G)) \rightarrow (\#F = \#G \equiv F \approx_D G)$$

The proof is in the Appendix.

(**776**) ⋆**Theorem:** Hume's Principle and Corollaries. Clearly, Hume's Principle fails to be modally strict: the left side of the principle has a truth value that doesn't vary from world to world – $\#F = \#G \rightarrow \Box(\#F = \#G)$ and $\#F \neq \#G \rightarrow \Box\#F \neq \#G$. But it is nevertheless derivable. We first note that (.1) the number of Gs numbers G:

(.1) $Numbers(\#G, G)$

From this we obtain Hume's Principle, i.e., (.2) the number of Fs is identical to the number of Gs if and only if F and G are equinumerous$_D$:

(.2) $\#F = \#G \equiv F \approx_D G$

Cf. Frege 1884 (§63–§72).[368] Note that the above version of Hume's Principle doesn't really differ from Frege's notwithstanding our introduction of the

[368] Interestingly, in 1893, Frege doesn't prove Hume's Principle as a biconditional (Tennant 2004, 108–9). Frege proves each direction as a separate theorem. The right-to-left direction is proved in §65, Theorem 32 [2013, 86], and the contraposed version of the left-to-right direction is proved in §69, Theorem 49 [2013, 93]. See May & Wehmeier 2019, for further discussion.

notion of discernibility into the analysis. Frege presumably thought that cardinal numbers could only count the discernible objects falling under a property, and our cardinal numbers will likewise count any discernible objects, ordinary or abstract, falling under a property. But our version of Hume's Principle is unlike Frege's in that $\#F$ has been rigidly defined in a modal setting.

If we expand the right side of Hume's Principle by definition (747.3) and apply theorem (750), we may conclude that (.3) the number of Fs is identical to the number of Gs if and only if there is a relation that maps$_D$ F onto G one-to-one:

(.3) $\#F = \#G \equiv \exists R(R\,|{:}\, F \xrightarrow[\text{onto}]{\text{1-1}}_D G)$

It is also a straightforward consequence of Hume's Principle and previous theorems that (.4) if F and G are materially equivalent with respect to the discernible objects, then the number of Fs is identical to the number of Gs:

(.4) $F \equiv_D G \rightarrow \#F = \#G$

(**777**) **Definition:** Natural Cardinals. Intuitively, cardinals are things that answer the question, "How many Fs are there?". Frege defined a cardinal number to be a number that belongs to some property. To adapt this suggestion to the present context, we define the notion of a *natural cardinal* as something that answers the question, "How many (discernible) Fs are there?". We say that x a *natural cardinal* if and only if, for some property G, x is identical to the number of Gs:

$$NaturalCardinal(x) \equiv_{df} \exists G(x = \#G)$$

Cf. Frege's definition of *Anzahl* in 1884, §72, and in 1893, §42 [2013, 58].[369] We

[369] In 1884, §72, Frege says:

> the expression
> "n is a Number"
> is to mean the same as the expression
> "there exists a concept such that n is the Number which belongs to it".
> Thus the concept of Number receives its definition, apparently, indeed, in terms of itself, but actually without any fallacy, since "the Number which belongs to the concept F" has already been defined.

We've captured Frege's definition in the present framework as $NaturalCardinal(x) \equiv_{df} \exists G(x = \#G)$. However, in 1893, §42 [2013, 58], Frege says:

> $\overset{\mathfrak{u}}{\smile}\; \mathfrak{N}\mathfrak{u} = \Gamma$ is the truth-value of: that there is a concept to which the cardinal number Γ belongs or, as we can also say, that Γ is a cardinal number. Therefore, we call the function $\overset{\mathfrak{u}}{\smile}\; \mathfrak{N}\mathfrak{u} = \xi$ the concept *cardinal number*.

So where we use # to operate on G to form $\#G$ (i.e., the number of Gs), Frege uses the symbol $\mathfrak{N}$ to operate on an extension $\mathfrak{u}$ of a concept to form $\mathfrak{N}\mathfrak{u}$, i.e., the number of the concept for which u is the extension.

now turn to a fact needed to show that that the definiens $\exists G(x = \#G)$ used in the definition above is equivalent, by a modally strict proof, to $\exists G(Numbers(x, G)$.

(**778**) **Theorem:** An Alternative Definiens of Natural Cardinal. Note first that (.1) if x numbers G, then x is a natural cardinal:

(.1) $Numbers(x, G) \rightarrow NaturalCardinal(x)$

With this fact, we may derive, by modally strict means, the equivalence of $\exists G(x = \#G)$ and $\exists G(Numbers(x, G))$:

(.2) $\exists G(x = \#G) \equiv \exists G(Numbers(x, G))$

This shows that, in the definition of $NaturalCardinal(x)$ (777), we could have used the condition $\exists G(Numbers(x, G))$ as the definiens. By contrast, as noted in (761), the biconditional $x = \#G \equiv Numbers(x, G)$ is a theorem but *not* a modally strict one.

(**779**) **Theorem:** Natural Cardinals and Modality. By modally strict means, it follows that (.1) if x is a natural cardinal, then necessarily x is a natural cardinal; and (.2) if x numbers G, then necessarily x is a natural cardinal:

(.1) $NaturalCardinal(x) \rightarrow \Box NaturalCardinal(x)$

(.2) $Numbers(x, G) \rightarrow \Box NaturalCardinal(x)$

Note how (.1) and (.2) contrast with (769.1).

(**780**) **Theorem:** Natural Cardinals Encode the Properties They Number. It is now provable that a natural cardinal x encodes a property F if and only if x just is the number of Fs:

$$NaturalCardinal(x) \rightarrow \forall F(xF \equiv x = \#F)$$

Clearly, once we apply the theory, then it follows from (.2) that if x is a natural cardinal and is equal to the number of planets (i.e., and $x = \#P$), then x encodes *being a planet*.

(**781**) **Lemma:** Fact About Non-Identity w.r.t. Discernibles. By (273.13), we know that *being a discernible object that is non-self-identical*, i.e., $[\lambda x\, D!x \,\&\, x \neq x]$, exists. Clearly, nothing whatsoever exemplifies this property:

$$\neg\exists y([\lambda x\, D!x \,\&\, x \neq x]y)$$

(**782**) **Definition** and **Theorem:** Zero. Given that $[\lambda x\, D!x \,\&\, x \neq x]\downarrow$ and that nothing exemplifies this property, we may define (.1) Zero to be the number of this property, i.e.,

(.1) $0 =_{df} \#[\lambda x\, D!x \,\&\, x \neq x]$

Using u as a (rigid) restricted variable ranging over discernible objects, we can write (.1) as follows:

$$0 =_{df} \#[\lambda u\, u \neq u]$$

Cf. Frege 1884, §74, and 1893, §41, Definition Θ [2013, 58].[370] By our theory of definitions and the fact that $[\lambda x\, D!x \,\&\, x \neq x]\downarrow$, it follows from the above definition that (.2) Zero exists:

(.2) $0\downarrow$

(**783**) **Theorem:** Zero is a Natural Cardinal.

$$NaturalCardinal(0)$$

We have taken the trouble to call out this theorem and the previous one, (782.2), because of their philosophical interest and significance. No mathematical axioms have been employed to achieve these results and all the mathematical notions required to derive them (e.g., $\#F$) have been defined in terms of non-mathematical notions.

(**784**) **Theorem:** Zero Numbers Empty Properties. As simple consequences of the previous theorems and definitions we have (.1) F fails to be exemplified by discernible objects iff Zero numbers F, and that (.2) F is exemplified by a discernible object if and only if some object other than Zero numbers it:

(.1) $\neg\exists u Fu \equiv Numbers(0,F)$

(.2) $\exists u Fu \equiv \exists x(Numbers(x,F) \,\&\, x \neq 0)$

Notice that since these theorems are modally strict, they hold necessarily. So, intuitively, at any possible world w, F is empty at w if and only if $Numbers(0,F)$. This shows that we've defined the natural cardinal Zero in such a way that it numbers empty properties at every world; there isn't a different number Zero at other worlds that number the empty properties there.

We also have the following facts: (.3) no discernible objects are actually F if and only if the number of Fs is Zero; (.4) if it is necessary that there are no discernible objects, then the number of Fs is Zero; and (.5) no discernible objects exemplify F at possible world w if and only if the number of the property *being F at w* is Zero:

(.3) $\neg\exists u \mathcal{A}Fu \equiv \#F = 0$

[370] We've kept the definition of Zero as close to Frege's definition as possible. But we could have used any other property that either actually or necessarily fails to be exemplified by discernible objects. For example, $\overline{L}$ (i.e., the negation of L, where L is the property $[\lambda x\, E!x \to E!x]$) is a property that is necessarily unexemplified and so necessarily unexemplified by discernible objects.

(.4) $\Box\neg\exists uFu \rightarrow \#F = 0$

(.5) $(w \models \neg\exists uFu) \equiv \#(F_w) = 0$

For the proof of (.5), recall the definition of F_w in (570.1) as $[\lambda x\, w \models Fx]$, and recall the fact (570.2) that $F_w\downarrow$. Kirchner observed (personal communication) that (.5) should convince one that the modally rigidified definition of *Numbers*(x, G) shows that the numbers arising at the actual world can number properties at other possible worlds; (.5) makes this explicit in the case of Zero.

(**785**) ⋆**Theorem:** Non-Strict Facts About Natural Cardinals. It is a theorem that (.1) a natural cardinal encodes all and only the properties it numbers:

(.1) $\mathit{NaturalCardinal}(x) \rightarrow \forall F(xF \equiv \mathit{Numbers}(x, F))$

Moreover, (.2) Zero encodes all and only the properties that no discernible object exemplifies:

(.2) $0F \equiv \neg\exists uFu$

Cf. Frege 1884 (§75). Finally, a property F is unexemplified by discernible objects if and only if the number of Fs is Zero:

(.3) $\neg\exists uFu \equiv \#F = 0$

Compare the left-to-right direction with Frege 1893, Theorem 97 (2013, 129).

14.4 Ancestrals and Relations on Discernibles

In this section, we look at the ancestrals of binary relations and then focus on the weak ancestrals of binary relations on discernibles. The definitions and theorems we study here will prepare us for the study of the predecessor relation in the next section.

(**786**) **Definition** and **Theorem:** Properties Hereditary w.r.t. a Binary Relation. Let us say that a property F *is hereditary with respect to* a binary relation G if and only if every pair of G-related objects are such that if the first exemplifies F then so does the second:

(.1) $\mathit{Hereditary}(F, G) \equiv_{df} F\downarrow \,\&\, G\downarrow \,\&\, \forall x\forall y(Gxy \rightarrow (Fx \rightarrow Fy))$

In what follows, we sometimes say F is G-hereditary instead of F is hereditary w.r.t. G. We now have the following theorem, in which we assert the existence of a relation defined in part in terms of (.1):

(.2) $[\lambda xy\, \forall F((\forall z(Gxz \rightarrow Fz) \,\&\, \mathit{Hereditary}(F, G)) \rightarrow Fy)]\downarrow$

This theorem establishes that the definiens of the following definition-by-= is significant.

(**787**) **Definition:** The Strong Ancestral of a Relation G. In 1879 (Part III, Proposition 76), 1884 (§79), and 1893 (§45, Definition K [2013, 60]), Frege defined the notion *y follows x in the G-series*. However later, Whitehead & Russell express this idea by saying *x is an ancestor of y with respect to relation G* (1910–1913, [1925–1927, ∗90, 549]). In the present work, we define the relation-forming operator *being a G-ancestor of*, written G^*, as: *being an x and y such that y exemplifies every property F that is both* (a) *exemplified by all the objects to which x is G-related and* (b) *hereditary with respect to G*:

$$G^* =_{df} [\lambda xy\, \forall F((\forall z(Gxz \to Fz) \,\&\, Hereditary(F,G)) \to Fy)]$$

In what follows, we sometimes refer to G^* as *the strong ancestral of* G. Clearly, given (786.2), (787), and our theory of definitions, we know $G^*\!\downarrow$.

(**788**) **Theorem:** Fundamental Fact about G^*. The following fundamental fact about the strong ancestral of G follows immediately by β-Conversion: x is a G-ancestor of y if and only if y exemplifies every property F that is both (a) exemplified by all the objects to which x is G-related and (b) hereditary with respect to G:

$$G^*xy \equiv \forall F((\forall z(Gxz \to Fz) \,\&\, Hereditary(F,G)) \to Fy)$$

If we speak mathematically for the moment (i.e., using mathematical notions we haven't defined within our system), we can think of any binary relation G as inducing sequences on its domain. That is, whenever Gab holds, there is a sense in which a comes immediately before b in a G-induced sequence. So, for example, if Gab, Gbc, Gcd, and Gbe, then we can say that G induces the sequence a,b,c,d (along with its subsequences) and it also induces the sequence a,b,e (along with its subsequences). We can understand the ancestral G^* as the relation that holds between the first member of a G-sequence and any later member of that sequence. Alternatively, one might say that the strong ancestral of G is the transitive closure of G.

(**789**) **Lemmas:** Facts About the Ancestral of G. The following are immediate consequences of (787) and (788): (.1) if x bears G to y, then x is a G-ancestor of y; (.2) if (a) x is an G-ancestor of y, (b) F is exemplified by every object to which x bears G, and (c) F is G-hereditary, then y exemplifies F; (.3) if (a) x exemplifies F, (b) x is an G-ancestor of y, and (c) F is G-hereditary, then y exemplifies F; (.4) if (a) x bears G to y and y is an G-ancestor of z, then x is an G-ancestor of z; (.5) if x is an G-ancestor of y, then something bears G to y; (.6) if x is an G-ancestor of y and y is a G-ancestor of z, then x is an G-ancestor of z; and (.8) if x is a G-ancestor of y, then x is a G-domain element:

(.1) $Gxy \to G^*xy$

(.2) $(G^*xy \,\&\, \forall z(Gxz \to Fz) \,\&\, Hereditary(F,G)) \to Fy$

(.3) $(Fx \,\&\, G^*xy \,\&\, Hereditary(F,G)) \to Fy$

(.4) $(Gxy \,\&\, G^*yz) \to G^*xz$

(.5) $G^*xy \to \exists zGzy$

(.6) $(G^*xy \,\&\, G^*yz) \to G^*xz)$

(.7) $G^*xy \to \exists zGxz$

Compare (.2) with Frege 1893, Theorem 123 [2013, 138] (though beware of Frege's use of italic a and Fraktur $\mathfrak{a}$ in the same formula), (.3) with Theorem 128 [2013, 139], (.4) with Theorem 129 [2013, 140], and (.5) with Theorem 124 [2013, 138]. Frege proved (.6) in 1879, Proposition 98.

(**790**) **Remark:** Digression on a Departure From Frege's Methods and Previous Methods. Those familiar with Frege's work would expect, at this point, a definition of the weak ancestral for an arbitrary binary relation G, i.e., expect a definition by identity for a new relation-term-forming operator G^+ ('the weak ancestral of G') using the expression *being an x and y such that either x is a strong G-ancestor of y or x is identical to y*. But if we were to offer such a definition, we wouldn't be able to guarantee that the definiens denotes a relation. That is, the λ-expression used as definiens in the following is not guaranteed to be significant:

$$G^+ =_{df} [\lambda xy\; G^*xy \vee x = y]$$

Of course, we might instead introduce a definition *by equivalence* that says: *x is a weak G-ancestor of y* ('$G^+(x,y)$') if and only if $G^*xy \vee x = y$ and explore theorems that govern this new formula. This is in fact the method used in Zalta 1999. In that work, it was asserted, as an axiom, that $Precedes(x,y)$ is a relation, and the weak ancestral of this relation was defined as the following condition:

$$Precedes^+(x,y) \equiv_{df} Precedes^*(x,y) \vee x = y$$

It was then asserted, as an axiom, that this condition also defined a relation. These axioms were then used to derive a general principle of induction and its corollary, the principle of mathematical induction.

But further investigation into the matter shows that the methods used in Zalta 1999 were stronger than strictly necessary and that we can achieve the goal of deriving the principle of mathematical induction using weaker methods. Indeed, the new axioms added in Zalta 1999 to derive the general principle of induction were not sufficiently restricted and require weakening, for with a little work, we can show that they lead to the McMichael/Boolos paradox.[371] To see the problem with the earlier work, suppose $\mathbb{P}$ is the *relation* of *predecessor*. Notice that if both the strong and weak ancestrals of $\mathbb{P}$ are also relations, written $\mathbb{P}^*$ and $\mathbb{P}^+$, then one can form the relation $[\lambda xy\, \mathbb{P}^+xy \,\&\, \neg\mathbb{P}^*xy]$, i.e., the relation of being an x and y such that x is a weak $\mathbb{P}$-ancestor of y but not a strong $\mathbb{P}$-ancestor of y. But this relation turns out to be equivalent to a unrestricted *relation of identity* given that the definitions of $\mathbb{P}$, $\mathbb{P}^*$, and $\mathbb{P}^+$ are defined generally.[372] An unrestricted relation of identity would reintroduce the McMichael/Boolos paradox, something that the present system currently avoids; see (192.3) and (192.4). So we have to take greater care when defining the weak ancestral of a relation G.

Fortunately, the error introduced by overlooking this consequence is not fatal. To address the problem, we'll focus on *relations on discernibles*, i.e., relations F such that $\Box\forall x\forall y(Fxy \rightarrow D!x \,\&\, D!y)$. This definition will allow us to introduce rigid restricted variables ranging over such relations, since any relation that satisfies the definition will necessarily do so. See (340), (341), and (792) below for further discussion of rigid restricted variables. Moreover, the strong ancestrals of relations on discernibles exist in virtue of (786.2) and (787), and the weak ancestrals of such relations can be defined in terms of identity w.r.t. discernibles, i.e., $=_D$. So, when G is a relation on discernibles, both its strong ancestral G^* and weak ancestral G^+ will exist and be well-defined.

Consequently, there is a there is a straightforward way to proceed. We first define and examine some theorems about relations on discernibles. Then we introduce rigid restricted variables $\underline{G}$ to range over such relations. We shall then define, for any relation on discernibles $\underline{G}$, its weak ancestral $\underline{G}^+$ in terms of its strong ancestral $\underline{G}^*$ and $=_D$. This method guarantees that the weak ancestrals of such relations are themselves relations. So when we eventually introduce the definition of *immediate predecessor* ($\mathbb{P}$) and show that $\mathbb{P}$ is a relation on discernibles, we'll then be able to form its weak ancestral $\mathbb{P}^+$, define *natural number*, and derive the Dedekind/Peano postulates more simply than was done in Zalta 1999. In particular, we shall no longer need to assert axiomatically that the weak ancestral of $\mathbb{P}$ is a relation. Moreover, this new method-

[371] Though Zalta 1999 included an Aczel model designed to show that the new axioms were consistent, in fact what it actually showed was that, if one gets the axioms right, one can consistently produce a Frege-style derivation of the natural numbers. This will be explained in some detail below.

[372] To see this, we can reason as follows:

$$\begin{array}{lcll} [\lambda xy\, \mathbb{P}^+xy \,\&\, \neg\mathbb{P}^*xy]zw & \equiv & \mathbb{P}^+zw \,\&\, \neg\mathbb{P}^*zw & \text{by } \beta\text{-Conversion} \\ & \equiv & (\mathbb{P}^*zw \vee z=w) \,\&\, \neg\mathbb{P}^*zw & \text{by definition of } \mathbb{P}^+zw \\ & \equiv & z=w & \text{by propositional logic} \end{array}$$

Hence the defined relation would hold between any two objects if and only if they are equal. There can be no such relation in object theory. This result could have been avoided if it had been asserted that the weak ancestral is a relation only relative to the domain of *Precedes*.

ology allows us to eliminate the modal axiom that was justified and added to object theory in Zalta 1999, when proving that every number has a successor. So the following investigation does not reprise Zalta 1999, but rather revises, corrects, and improves on the presentation there.

(**791**) **Definition and Theorem:** Relations on Discernibles. When $n \geq 1$, we say that (.1) F is a (n-ary) *relation on discernibles* just in case F exists and, necessarily, any objects $x_1, \ldots, x_n$ that exemplify F are all discernible objects:

(.1) $OnDiscernibles(F) \equiv_{df} F\downarrow \; \& \; \Box\forall x_1 \ldots \forall x_n(Fx_1 \ldots x_n \rightarrow (D!x_1 \; \& \; \ldots \; \& \; D!x_n)) \; (n \geq 1)$

By simple, modally strict reasoning from (.1) and the 4 schema, it follows that (.2) for any relation F, if F is a relation on discernibles, then necessarily, F is a relation on discernibles:

(.2) $\forall F(OnDiscernibles(F) \rightarrow \Box OnDiscernibles(F))$

Hence $OnDiscernibles(F)$ is a rigid condition on relations.

(**792**) **Remark:** Rigid Restricted Variables for Relations on Discernibles. Our goal is now to establish that $OnDiscernibles(F)$ is a rigid restriction condition on the variable F. We therefore have to show that it meets the conditions defined in (336) for a restriction condition as well as the additional condition laid down in (340) for rigid restriction conditions. Since the condition has a single free variable, we need only show (i) that the condition is strictly non-empty, i.e., that $\vdash \exists F(OnDiscernibles(F))$, and (ii) that it has existential import, i.e., that $\vdash OnDiscernibles(\Pi) \rightarrow \Pi\downarrow$, for any relation term Π.

Clearly, it is provable that there are relations on discernibles. $=_E$ and $=_D$ are examples of such. For the purposes of this discussion, we show this only for $=_E$ and leave the latter as an exercise. $=_E$ is a relation on discernibles since it is provable that $\Box\forall x\forall y(x =_E y \rightarrow O!x \; \& \; O!y)$ and since $O!z \rightarrow D!z$, it follows that $\Box\forall x\forall y(x =_E y \rightarrow D!x \; \& \; D!y)$. Analogous reasoning shows that $=_D$ is a relation on discernibles.

Moreover, it should be clear that $OnDiscernibles(F)$ has existential import, since that is the first conjunct of the definition (791.1): if a relation term satisfies the condition, then that term denotes.

These observations, together with the fact that the universal generalization of (791.2) is a modally strict theorem, establish that $OnDiscernibles(F)$ is itself a rigid restriction condition on relations, as defined in (340). We may therefore introduce the following *rigid* restricted variables:

$\underline{F}, \underline{G}, \underline{H}, \ldots$ range over relations on discernibles

(**793**) **Definition:** The Weak Ancestral of Relations on Discernibles. If $\underline{G}$ is any relation on discernibles and $=_D$ is the relation of identity with respect to the discernible objects (273.17), we define the relation *being a weak $\underline{G}$ ancestor of* (written $\underline{G}^+$) as the relation *being x and y such that x is an $\underline{G}$ ancestor of y or x is identical$_D$ to y*:

$$\underline{G}^+ =_{df} [\lambda xy\, \underline{G}^* xy \lor x =_D y]$$

Cf. Frege 1879, §29; 1893, §46, Definition Λ [2013, 60]. We leave it as an exercise to show that the definiens is provably significant. So the identity $\underline{G}^+ = [\lambda xy\, \underline{G}^* xy \lor x =_D y]$ is a theorem and $\underline{G}^+$ provably exists.

(**794**) **Theorem:** Basic Fact About the Weak Ancestral of Relations on Discernibles. It follows immediately by β-Conversion that x bears the weak ancestral of $\underline{G}$ to y whenever x bears the strong ancestral of $\underline{G}$ to y or x is *identical$_D$* to y:

$$\underline{G}^+ xy \equiv \underline{G}^* xy \lor x =_D y$$

We shall sometimes use Frege's reading of the left condition as: *y is a member of the $\underline{G}$ series beginning with x* (1884, §81; 1893, §46 [2013, 60]). Though Frege's reading has a slightly increased cognitive load (the variable for the object occurring later in the $\underline{G}$ series occurs earlier in the expression), the effect is mitigated by the phrase *beginning with x*. The intuition here is that $\underline{G}$ induces sequences on its domain.

(**795**) **Lemmas:** Facts About the Weak Ancestral of $\underline{G}$. The following are immediate consequences of our definitions: (.1) if x bears $\underline{G}$ to y, then x is a weak $\underline{G}$-ancestor of y; (.2) if (a) x exemplifies F, (b) x is a weak $\underline{G}$-ancestor of y, and (c) F is $\underline{G}$-hereditary, then y exemplifies F; (.3) if x is a weak $\underline{G}$-ancestor of y and y bears $\underline{G}$ to z, then x is a $\underline{G}$-ancestor of z; (.4) if x is a $\underline{G}$-ancestor of y and y bears $\underline{G}$ to z, then x is a weak $\underline{G}$-ancestor of z; (.5) if x bears $\underline{G}$ to y, and y is a weak $\underline{G}$-ancestor of z, then x is a $\underline{G}$-ancestor of z; (.6) if x is a weak $\underline{G}$-ancestor of y and y is a weak $\underline{G}$-ancestor of z, then x is a weak $\underline{G}$-ancestor of z; and (.7) if x is a $\underline{G}$-ancestor of y, then x is a weak $\underline{G}$-ancestor of something that bears $\underline{G}$ to y:

(.1) $\underline{G}xy \to \underline{G}^+ xy$

(.2) $(Fx \,\&\, \underline{G}^+ xy \,\&\, Hereditary(F, \underline{G})) \to Fy$

(.3) $(\underline{G}^+ xy \,\&\, \underline{G}yz) \to \underline{G}^* xz$

(.4) $(\underline{G}^* xy \,\&\, \underline{G}yz) \to \underline{G}^+ xz$

(.5) $(\underline{G}xy \,\&\, \underline{G}^+ yz) \to \underline{G}^* xz$

(.6) $(\underline{G}^+ xy \,\&\, \underline{G}^+ yz) \to \underline{G}^+ xz$

(.7) $\underline{G}^* xy \to \exists z(\underline{G}^+ xz \,\&\, \underline{G}zy)$

Cf. (.2) with Frege 1893, Theorem 144 [2013, 143]; (.3) with Theorem 134 [2013, 142]; (.5) with Theorem 132 [2013, 140]; and (.7) with Theorem 141 [2013, 143].[373] Frege proves these principles for weak ancestrals for relations generally, however all of his relations were in fact relations on discernibles and so the restriction to relations on discernibles is not an egregious one. We haven't discovered a statement or proof of (.6) (i.e., transitivity of the weak ancestral) in Frege 1879 or 1893.

Note, finally, that $\underline{G}^{+}xy \rightarrow \exists z\underline{G}xz$ is not a theorem. An empty relation is a relation on discernibles, by failure of antecedents. Consider such an empty relation, $\underline{G}$, and let a be a discernible object. Then since $a =_D a$, it follows from (794) that $\underline{G}^{+}aa$ holds. But since $\underline{G}$ is empty, $\neg\exists y(Gay)$.

(**796**) **Definitions and Theorems:** One-to-One Relations and One-to-One Relations on Discernibles and Their Ancestrals. We say that (.1) a binary relation G is *one-to-one*, written *1-1*(G), just in case for any objects x, y, and z, if x bears G to z and y bears G to z, then x is identical to y:

(.1) $\mathit{1\text{-}1}(G) \ \equiv_{df}\ G{\downarrow}\ \&\ \forall x\forall y\forall z(Gxz\ \&\ Gyz \rightarrow x = y)$

We then have the following facts about relations on discernibles that are also one-to-one: (.2) if $\underline{G}$ is a one-to-one relation on discernibles, then if x bears $\underline{G}$ to y and z is a $\underline{G}$-ancestor of y, then z is a weak $\underline{G}$-ancestor of x; (.3) if $\underline{G}$ is a one-to-one relation on discernibles, then if x bears $\underline{G}$ to y and x fails to be a $\underline{G}$-ancestor of x, then y fails to be a $\underline{G}$-ancestor of y; (.4) if $\underline{G}$ is a one-to-one relation on discernibles then if x fails to be a $\underline{G}$-ancestor of x and x is a weak $\underline{G}$-ancestor of y, then y fails to be a weak $\underline{G}$-ancestor of y:

(.2) $\mathit{1\text{-}1}(\underline{G}) \rightarrow ((\underline{G}xy\ \&\ \underline{G}^{*}zy) \rightarrow \underline{G}^{+}zx)$

(.3) $\mathit{1\text{-}1}(\underline{G}) \rightarrow ((\underline{G}xy\ \&\ \neg\underline{G}^{*}xx) \rightarrow \neg\underline{G}^{*}yy)$

(.4) $\mathit{1\text{-}1}(\underline{G}) \rightarrow ((\neg\underline{G}^{*}xx\ \&\ \underline{G}^{+}xy) \rightarrow \neg\underline{G}^{*}yy)$

Some of these facts are used a bit later, to establish that no natural number is less than itself or is a predecessor of itself. But we won't need them to prove the following, important principle.

(**797**) **Theorem:** Generalized Induction. Let $\underline{G}$ be any relation on discernibles. Then it is a theorem that if (a) z exemplifies F and (b) any two objects x and y

[373] The version of this theorem asserted in Zalta 1999 was incorrectly asserted — the 'proof' incorrectly referenced an ill-formed λ-expression that contained the non-propositional formula $G^{+}(x,z)$, which violated the formation rules of Zalta 1999. Fortunately, the version of (.7) asserted in Zalta 1999 wasn't used in that paper to derive any of the Dedekind/Peano postulates as theorems. Of course, the issue here is overshadowed by the issue discussed in Remark (790). Nevertheless, it is important to point out the error, and to note that in its present formulation, (.7) is restricted to relations on discernibles and is indeed a theorem that will be used below in the proof of an important fact, namely (796.2).

having z as a weak $\underline{G}$ ancestor are such that, if x bears $\underline{G}$ to y then Fx implies Fy, then (c) every object of which z is a weak $\underline{G}$ ancestor exemplifies F. Formally:

Generalized Induction
$$[Fz \,\&\, \forall x \forall y((\underline{G}^{+}zx \,\&\, \underline{G}^{+}zy) \rightarrow (\underline{G}xy \rightarrow (Fx \rightarrow Fy)))] \rightarrow \forall x(\underline{G}^{+}zx \rightarrow Fx)$$

We can state this theorem a bit more intuitively if we informally introduce a variation on the notion F is hereditary with respect to $\underline{G}$. Let us use the expression:

F is hereditary on the $\underline{G}$-series beginning with z

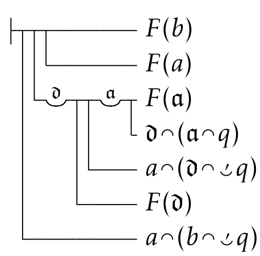
whenever:

$$\forall x \forall y((\underline{G}^{+}zx \,\&\, \underline{G}^{+}zy) \rightarrow (\underline{G}xy \rightarrow (Fx \rightarrow Fy)))$$

Then we may state **Generalized Induction** as follows: if both (i) z exemplifies F and (ii) F is hereditary on the $\underline{G}$-series beginning with z, then every member of that series beginning with z exemplifies F.

This is a variant of Frege 1893, Theorem 152 [2013, 148]. Those interested in the differences between the above version and Frege's may find the following Remark useful.

(**798**) **Remark:** Digression on the Changes to Frege's Version of Generalized Induction. If we put aside the fact that we've restricted Generalized Induction to relations on discernibles, then (797) differs from Frege's Theorem 152 in two ways: (a) the consequent of the principle is expressed as a universal claim, and (b) an additional, though strictly unnecessary, conjunct has been added to the second conjunct of the antecedent. In this Remark, we explain the differences in more detail and justify both changes. Consider Frege's Theorem 152:



This can be rewritten in our notation for predication, conditionals, and quantification as follows, where (i) a and b continue to serve, as in the above, as individual variables, (ii) the Gothic letters $\mathfrak{d}$ and $\mathfrak{a}$ are replaced by the variables x and y, respectively, (iii) q is replaced by G (which for the purposes of this Remark, can be any relation), and (iv) $\smile q$ is replaced by G^{+}:

$$[G^{+}ab \rightarrow \forall x(Fx \rightarrow (G^{+}ax \rightarrow \forall y(Gxy \rightarrow Fy)))] \rightarrow (Fa \rightarrow Fb)$$

If we validly swap G^+ab (i.e., the antecedent of the antecedent) with Fa (i.e., the antecedent of the consequent) and apply Importation (88.7.b) in a couple of places, this starts to look more familiar:

$$[Fa \,\&\, \forall x(Fx \,\&\, G^+ax \to \forall y(Gxy \to Fy))] \to (G^+ab \to Fb)$$

In other words, if a exemplifies F and every F-object in the G-series beginning with a passes F to everything to which it is G-related, then if b is in the G-series beginning with a, b exemplifies F.

Now if we rearrange the second conjunct of the antecedent a bit more using (99.7), symmetry of &, and Exportation (88.7.a), the last claim displayed above is equivalent to:

$$[Fa \,\&\, \forall x\forall y(G^+ax \to (Gxy \to (Fx \to Fy))] \to (G^+ab \to Fb)$$

Since b is being used here as a free variable, GEN tells us if the above is derivable, then it holds for every b. So by classical quantification theory (95.2):

$$\text{(A)}\quad [Fa \,\&\, \forall x\forall y(G^+ax \to (Gxy \to (Fx \to Fy))] \to \forall b(G^+ab \to Fb)$$

Now compare (A) with (797): if you replace G by $\underline{G}$, a by z, and b by x in (A), then (A) becomes similar to (797) with the only difference being that (A) eliminates G^+ay from the second conjunct of the antecedent of (797). Frege, of course, recognized that G^+ay isn't needed in the statement of the theorem, for it is provable in his system that if x is in the G-series beginning with z and x bears G to y, then y is in the G-series beginning with z, i.e., it is a theorem that $(G^+zx \,\&\, Gxy) \to G^+zy$.[374] So (797) and (A) are equivalent, modulo the former's restriction to relations on discernibles.

But we've formulated General Induction as in (797) because once it is instantiated in the manner described in (812) (i.e., by instantiating z to Zero and $\underline{G}$ to $\mathbb{P}$), one can rewrite the result using rigid restricted variables over the natural numbers to obtain the simple, classical statement of the Principle of Mathematical Induction. See the statement and proof of (812) below.

14.5 Predecessor

(**799**) **Remark:** Frege's Definition. In 1884 (§76) and 1893 (§43, Definition H [2013, 58]), Frege defined *immediately precedes* as follows:

$$\mathit{Precedes}(x,y) \equiv_{df} \exists F\exists z(Fz \,\&\, y=\#F \,\&\, x=\#F^{-z}) \qquad \text{(H)}$$

[374] In our system, we can prove a version for relations on discernibles. Assume $\underline{G}^+zx \,\&\, \underline{G}xy$. Then by (795.3), $\underline{G}^*zy$. Hence, by $\vee$I and definition (794), $\underline{G}^+zy$.

For those who have never encountered Frege's definition before, the insight can be made clear by way of an intuitive example. Though we haven't yet defined '1' and '2', we pre-theoretically know that *Precedes*$(1,2)$. So given Frege's definition, the following should hold: $\exists F \exists z(Fz \,\&\, 2 = \#F \,\&\, 1 = \#F^{-z})$. This latter condition is intuitively true if we let F be the property *author of Principia Mathematica* and let r be Bertrand Russell. Then we know Fr, since Russell is an author of *Principia Mathematica*. Moreover, we know $2 = \#F$, since 2 is the number of the property *author of Principia Mathematica* — both Russell and Whitehead coauthored the work. Finally, we know $1 = \#F^{-r}$, since 1 is the number of the property *author of Principia Mathematica not identical to Russell* — there is only one object, namely, Whitehead, that exemplifies this property. So, conjoining Fr, $2 = \#F$, and $1 = \#F^{-r}$, and then existentially generalizing on r and F, we have the definiens when (H) is instantiated to 1 and 2, i.e., we have the definiens of *Precedes*$(1,2)$.

In the next item (800), we shall modify Frege's methodology somewhat. Instead of the definiens for (H), we shall use the formula:

$$\exists F \exists u(Fu \,\&\, \mathit{Numbers}(y,F) \,\&\, \mathit{Numbers}(x,F^{-u})) \qquad (\mathrm{H}')$$

Note that (H′) is formulable entirely in terms of non-mathematical primitives, since *Numbers*(y,F) and *Numbers*(x,F^{-u}) are eliminable in terms of primitive notation. Moreover we'll first assert, as an *axiom*, that (H′) (which is a condition on x and y) defines a relation, and then *define* the mathematical notion of *predecessor*, which we'll symbolize as $\mathbb{P}$, in terms of this relation.

So (H′), and thus our axiom, alters Frege's definition of predecessor in two ways. First, since our natural cardinals only count the discernible objects exemplifying a property F, the existential quantifier $\exists z$ in (H) is replaced in (H′) by the restricted quantifier $\exists u$ (with u replacing z throughout the remainder of the formula), where u is a rigid restricted variable ranging over discernible objects. Given this replacement, $\#F^{-u}$ is already defined in (752.2) as the number of the property *being an F-object not identical to u*.

Second, (H′) also differs from (H) by the substitution of *Numbers*(y,F) for $y = \#F$ and *Numbers*(x,F^{-u}) for $x = \#F^{-z}$. Whereas Frege was not concerned about modal contexts, we want *predecessor* to operate in any modal context. So whereas the condition $y = \#F$ identifies y as the object that *in fact* numbers F, the condition *Numbers*(y,F) identifies y, no matter what the modal context, as the object that numbers the Fs in that context. In non-modal contexts, these two conditions are equivalent, since it follows immediately from (776.1)$\star$ that *Numbers*$(y,F) \equiv y = \#F$. Moreover, the condition *Numbers*(y,F) is preferable because it is several steps closer to primitive notation than the condition $y = \#F$; it makes no use of definite descriptions and doesn't involve an identity formula (which is defined in object theory, unlike in Frege's theory, where identity is a primitive).

(**800**) **Axiom:** The Existence of an Ordering Relation. We now assert that the following exists (as a relation): *being an x and y such that, for some property F and some discernible object u that exemplifies F, y numbers F and x numbers the F*s *other than u*:

$$[\lambda xy\, \exists F \exists u(Fu \,\&\, Numbers(y,F) \,\&\, Numbers(x,F^{-u}))]\downarrow$$

It should be emphasized that this relation is a non-mathematical one. No mathematical primitives are used to express this relation and, as with other ordering relations, it can be expressed entirely in logical terms. For example, a partial order is any relation G that is irreflexive, transitive, and anti-symmetric, where these conditions can be defined in the usual manner, without any mathematics. So the above relation simply orders those abstract objects that, intuitively, encode equivalence classes of properties whose discernible exemplifiers can be put in one-to-one correspondence. There is nothing inherently mathematical in this relation, just as there is nothing inherently mathematical about the relations *x is to the left of y*, *x is on top of y*, *x is outside of y*, *x is after y*, etc.

(**801**) **Definition** and **Theorems:** Predecessor. We now introduce the new binary relation constant $\mathbb{P}$ for the relation *immediate predecessor*:

(.1) $\mathbb{P} =_{df} [\lambda xy\, \exists F \exists u(Fu \,\&\, Numbers(y,F) \,\&\, Numbers(x,F^{-u}))]$

This definition-by-identity now yields, by our theory of definitions, the claim $\mathbb{P} = [\lambda xy\, \exists F \exists u(Fu \,\&\, Numbers(y,F) \,\&\, Numbers(x,F^{-u}))]$ and hence that the relation *immediate predecessor* exists:

(.2) $\mathbb{P}\downarrow$

It also follows by β-Conversion that (.3) *x immediately precedes y* if and only if there is a property F and a discernible object u such that (a) u exemplifies F, (b) y numbers F, and (c) x numbers *being-F-but-not-identical*$_D$*-to-u*:

(.3) $\mathbb{P}xy \equiv \exists F \exists u(Fu \,\&\, Numbers(y,F) \,\&\, Numbers(x,F^{-u}))$

This theorem yields necessary and sufficient conditions for the predecessor relation.

Since (.3) is a modally strict theorem, a Rule of Substitution (160.2) allows us to substitute the left and right sides for one another in any formula where one occurs as a subformula.

(**802**) **Theorem:** Predecessor is a Rigid, One-to-One, and Functional Relation. We'll later show that *predecessor* is a relation on discernibles, but for now we establish that (.1) if x precedes y, it necessarily does; (.2) *predecessor* is a rigid relation; (.3) *predecessor* is a one-to-one relation; and (.4) *predecessor* is a functional relation:

(.1) $\mathbb{P}xy \rightarrow \Box\mathbb{P}xy$

(.2) $Rigid(\mathbb{P})$

(.3) $1\text{-}1(\mathbb{P})$

(.4) $\mathbb{P}xy \,\&\, \mathbb{P}xz \rightarrow y = z$

In connection with (.3), cf. Frege 1884, §78; and 1893, Theorem 89 [2013, 127]. In connection with (.4), cf. Frege 1893, Theorem 71 [2013, 113].

(**803**) **Theorems:** Predecessor is a Relation on Discernibles. We now show work our way to the principle that predecessor is a relation on discernibles. To establish this, we prove (.1) there are objects x and y such that x *precedes* y; (.2) if x is a natural cardinal other than Zero, there is a object that precedes x; (.3) natural cardinals are discernible; (.4) if x precedes y, both x and y are natural cardinals; and (.5) if x precedes y, x and y are both discernible:

(.1) $\exists x \exists y \mathbb{P}xy$

(.2) $NaturalCardinal(x) \,\&\, x \neq 0 \rightarrow \exists y \mathbb{P}yx$

(.3) $NaturalCardinal(x) \rightarrow D!x$

(.4) $\mathbb{P}xy \rightarrow (NaturalCardinal(x) \,\&\, NaturalCardinal(y))$

(.5) $\mathbb{P}xy \rightarrow (D!x \,\&\, D!y)$

Finally, it may be useful to know that *being an x that numbers F* exists:

(.6) $[\lambda x\, Numbers(x, F)]\downarrow$

Given that (.5) is a modally strict theorem, it follows by GEN and RN that $\mathbb{P}$ is a relation on discernibles (791.1). So we can now instantiate $\mathbb{P}$ into any universal claim governing such relations.

(**804**) **Theorems:** Strong Ancestral of Predecessor. Since (a) $\mathbb{P}\downarrow$, and (b) the definition in (787) yields a universal identity claim that is general with respect to every binary relation G, we can instantiate that universal claim to $\mathbb{P}$ to obtain, as a theorem, that (.1) *the strong ancestral of predecessor* is (identical to) the relation: *being objects x and y such that y exemplifies every F such that is* (a) *exemplified by every object that x precedes and* (b) *hereditary w.r.t. predecessor*:

(.1) $\mathbb{P}^* = [\lambda xy\, \forall F((\forall z(\mathbb{P}xz \rightarrow Fz) \,\&\, Hereditary(F, \mathbb{P})) \rightarrow Fy)]$

Hence, by (107.2), we know (.2) the strong ancestral of predecessor exists:

(.2) $\mathbb{P}^*\downarrow$

And it also follows by β-Conversion and the definition of $Hereditary(F,G)$ (786) that (.3) x bears the strong ancestral of predecessor to y just in case y exemplifies every property F that is both (a) exemplified by all the objects of which x is a predecessor and (b) hereditary with respect to predecessor:

(.3) $\mathbb{P}^*xy \equiv \forall F((\forall z(\mathbb{P}xz \to Fz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Fx' \to Fy'))) \to Fy)$

It is now worth a moment's reflection to consider the fact that the theorems in (789) governing the strong ancestral of a relation now apply when $\mathbb{P}$ is substituted for G. Finally, it helps to prove that (.4) the strong ancestral of predecessor is a rigid relation:

(.4) $Rigid(\mathbb{P}^*)$

This fact plays a role when we reprise, in our modal setting, the Fregean derivation of the claim that every number has a successor (818). In the Appendix, we give two proofs of (818); the second one is more similar to Frege's in that it appeals to a theorem concerning $\#F$. Thus, (.4) is used in the proof that the weak ancestral of $\mathbb{P}$ is rigid (806.3), which is used in turn in (817.3) and thus plays a role in the proof of the lemma that n immediately precedes the number of the property *bearing the weak ancestral of predecessor to n* (817.7). This last lemma quickly yields that every number has a successor (818). However, our first proof of (818) in the Appendix doesn't require (.4). In that proof, we appeal to the defined notion $Numbers(x,F)$ instead of $\#F$; the proof of (.4) won't be needed, for example, in the proof of (817.2) or (817.6).

(**805**) **Lemma:** Zero, Predecessor, and its Ancestrals. It is a consequence of our definitions that nothing is an immediate predecessor of Zero:

(.1) $\neg\exists x\mathbb{P}x0$

Cf. Frege 1893, Theorem 108 [2013, 131]. It follows *a fortiori* that no natural cardinal is an immediate predecessor of Zero.

Moreover, it also follows that (.2) nothing is a predecessor ancestor of Zero:

(.2) $\neg\exists x\mathbb{P}^*x0$

Cf. Frege 1893, Theorem 126 [2013, 138]. Clearly, then, (.3) Zero is not a predecessor ancestor of itself:

(.3) $\neg\mathbb{P}^*00$

(**806**) **Theorems:** The Weak Ancestral of Predecessor. Since $\mathbb{P}$ exists and is provbably a relation on discernibles, the weak ancestral of $\mathbb{P}$ is defined by an instance of (793). So we know that (.1) $\mathbb{P}^+$ exists:

(.1) $\mathbb{P}^+\!\downarrow$

It immediately follows by β-Conversion that (.2) x is a weak *predecessor*-ancestor of y if and only if either x is a predecessor ancestor of y or x is identical$_D$ to y:

(.2) $\mathbb{P}^{+}xy \equiv \mathbb{P}^{*}xy \vee x =_D y$

With (.2), we've reached a position where we can define the remaining primitive notions, and derive the postulates, of Dedekind/Peano number theory. But it is worth pausing for a moment to reflect on the fact that when $\mathbb{P}$ is substituted for $\underline{G}$ in the theorems in (795) – (797), we obtain specific new theorems about $\mathbb{P}$ and its ancestrals. It is also worth noting that (.3) the weak ancestral of predecessor is a rigid relation:

(.3) $Rigid(\mathbb{P}^{+})$

(.3) will be relevant for the proof of lemma (817.3), which in turn is used in the key lemma (817.7) needed for the Fregean-style proof of the claim that every natural number has a successor.

14.6 Deriving the Dedekind/Peano Postulates

(**807**) **Definition** and **Theorems:** Natural Numbers. Since $[\lambda x\ \mathbb{P}^{+}0x]$ clearly exists, we may now define (.1) *being a natural number* as *being an x such that Zero bears the the weak ancestral of predecessor to x*:

(.1) $\mathbb{N} =_{df} [\lambda x\, \mathbb{P}^{+}0x]$

Thus, to be a natural number is, in Frege's terminology, to be a member of the predecessor-series beginning with Zero. Frege calls such numbers *finite cardinal* numbers (1893, §108 [2013, 137]). It now follows, in the usual way, that (.2) the property *being a natural number* exists; and (.3) x is a natural number if and only if Zero bears the weak ancestral of *predecessor* to x:

(.2) $\mathbb{N}\downarrow$

(.3) $\mathbb{N}x \equiv \mathbb{P}^{+}0x$

(**808**) **Theorem:** Zero is a Natural Number.

$\mathbb{N}0$

In the present system, the proof of this claim is trivial and with it we have derived the first Dedekind/Peano postulate.

Interestingly, Frege (1893) doesn't seem to prove this claim as a theorem, possibly because, in his system, it is a trivial consequence of definitions and facts about identity. He proves only the general theorem $G^{+}xx$, i.e., that x is a weak G-ancestor of itself (1893, Theorem 140 [2013, 143]), and doesn't label the instance $\mathbb{P}^{+}00$ as a separate theorem.

(**809**) **Lemmas:** $\mathbb{N}$ is a Rigid Property and $\mathbb{N}x$ is a Rigid Restriction Condition. The following lemmas will prove to be important. (.1) if x is a number, then necessarily x is a number; (.2) *being a number* is a rigid property:

(.1) $\mathbb{N}x \to \Box\mathbb{N}x$

(.2) $Rigid(\mathbb{N})$

Clearly, then, the formula $\mathbb{N}x$ is a rigid restriction condition, as this was defined in (340): it has a single free variable; it follows from (808) that it is strictly non-empty (i.e., there is a modally strict proof that $\exists x\mathbb{N}x$), it has strict existential import since it is easy to give a modally strict proof for the claim $\mathbb{N}\kappa \to \kappa\downarrow$, for any individual term κ; and by applying GEN to (.1), we obtain a modally strict proof of $\forall x(\mathbb{N}x \to \Box\mathbb{N}x)$. Consequently, we may introduce rigid restricted variables to range over numbers. We use m, n, k, and if needed, i and j, as such variables.

(**810**) **Theorem:** Zero Is Not the Successor of Any Natural Number. It also follows that no natural number is an immediate predecessor of Zero.

$$\neg\exists n\mathbb{P}n0$$

Cf. Frege, 1893, Theorem 126 [2013, 138]. With this theorem, we have derived the second Dedekind/Peano postulate.

(**811**) **Theorems:** No Two Natural Numbers Have the Same Successor. From the fact that *predecessor* is a one-to-one relation generally (802.3), it follows *a fortiori* that it is one-to-one with respect to the natural numbers. Hence, no two natural numbers have the same successor, i.e.,

$$\forall n\forall m\forall k(\mathbb{P}nk \,\&\, \mathbb{P}mk \to n = m)$$

With (811), we have derived the third Dedekind/Peano postulate. Before we derive the fourth postulate, that every number has a successor, we first derive the fifth postulate, namely, the principle of mathematical induction.

(**812**) **Theorem:** Mathematical Induction. Since $\mathbb{P}$ is a relation on discernibles, the Principle of Mathematical Induction falls out as a corollary to (797). For every property F, if Zero exemplifies F and Fn implies Fm whenever n and m are any two successive natural numbers, then every natural number exemplifies F:

$$\forall F[F0 \,\&\, \forall n\forall m(\mathbb{P}nm \to (Fn \to Fm)) \to \forall nFn]$$

With this theorem, we have derived the fifth Dedekind-Peano postulate. We now work our way towards a proof of the fourth postulate, that every natural number has a unique successor (818).

(**813**) **Lemma:** Natural Numbers are Natural Cardinals. It is a consequence of our definitions that (.1) natural numbers are natural cardinals, and (.2) discernible:

(.1) $\mathbb{N}x \to \mathit{NaturalCardinal}(x)$

(.2) $\mathbb{N}x \to D!x$

These lemmas play a role in the proof that every natural number has a successor.

(**814**) **Lemma:** Successors of Natural Numbers are Natural Numbers. A successor of a natural number is a natural number:

(.1) $\mathbb{P}nx \to \mathbb{N}x$

Recall that this principle is listed by some as one of the basic postulates of number theory; see postulate 6 of Remark (740).

Moreover, it is also a theorem that (.2) if a natural number is a predecessor ancestor of x, then x is a natural number; and (.3) if a natural number is a weak predecessor ancestor of x, then x is a natural number:

(.2) $\mathbb{P}^{*}nx \to \mathbb{N}x$

(.3) $\mathbb{P}^{+}nx \to \mathbb{N}x$

(**815**) **Lemma:** Predecessors of Natural Numbers are Natural Numbers. A predecessor of a natural number is a natural number:

$$\mathbb{P}xn \to \mathbb{N}x$$

(**816**) **Lemma:** Predecessor is a Functional Relation on the Natural Numbers. Given that predecessor is a functional relation *tout court*, it is a functional relation when restricted to the natural numbers:

$$\mathbb{P}nm \,\&\, \mathbb{P}nk \to m = k$$

Recall that this theorem is counted, by some, as one of the basic postulates of number theory; see postulate 7 of Remark (740).

(**817**) **Lemmas:** Key Facts About the Ancestrals of Predecessor and Natural Numbers. We now prove four lemmas that are used in the proof that every number has a successor, namely, (.1) no natural number is a predecessor ancestor of itself; (.2) if an object y immediately precedes a natural number x, then an object z numbers *being a weak-predecessor ancestor of y* if and only if z numbers *being a weak-predecessor ancestor of x other than x*; (.3) if an object y immediately precedes a natural number x, then the number of *being a weak-predecessor ancestor of y* is identical to the number of *being a weak-predecessor ancestor of x other than x*; (.4) *being an object x that immediately precedes something that numbers being a weak predecessor-ancestor of x* exists; (.5) *being an object x that precedes the number of the concept: being a weak predecessor-ancestor of x* exists; (.6) every natural number n immediately precedes an object that numbers *being a weak predecessor-ancestor of n*; (.7) every natural number n immediately precedes the number of *being a weak predecessor-ancestor of n*:

(.1) $\forall x(\mathbb{N}x \rightarrow \neg\mathbb{P}^{*}xx)$

(.2) $(\mathbb{N}x \,\&\, \mathbb{P}yx) \rightarrow (Numbers(z, [\lambda z\, \mathbb{P}^{+}zy]) \equiv Numbers(z, [\lambda z\, \mathbb{P}^{+}zx]^{-x}))$

(.3) $(\mathbb{N}x \,\&\, \mathbb{P}yx) \rightarrow \#[\lambda z\, \mathbb{P}^{+}zy] = \#[\lambda z\, \mathbb{P}^{+}zx]^{-x}$

(.4) $[\lambda x\, \exists y(Numbers(y, [\lambda z\, \mathbb{P}^{+}zx]) \,\&\, \mathbb{P}xy)]\downarrow$

(.5) $[\lambda x\, \mathbb{P}x\#[\lambda z\, \mathbb{P}^{+}zx]]\downarrow$

(.6) $\forall n \exists y(Numbers(y, [\lambda z\, \mathbb{P}^{+}zn]) \,\&\, \mathbb{P}ny)$

(.7) $\forall n \mathbb{P}n\#[\lambda z\, \mathbb{P}^{+}zn]$

Lemma (.3) corresponds to Frege's Theorem 149 (1893 [2013, 147]). In the Appendix, we give two proofs of (.7). The first derives it as an easy consequence of (.6). But, as an alternative, we give the 'Fregean' proof of (.7), by induction. In that induction, the base case corresponds to Frege's Theorem 154 (1893 [2013, 147]), while the inductive case corresponds to Frege's Theorem 155 (1893 [2013, 149]).

(**818**) **Theorem:** Every Natural Number Has a Unique Natural Number Successor. We can now prove that every natural number n, there is a unique natural number m that n precedes:

$$\forall n \exists! m \mathbb{P}nm$$

With this theorem, we have derived the last of the 5 main Dedekind-Peano postulates. Later in this chapter, we complete the derivation of second-order Peano Arithmetic (PA2) by (a) introducing the recursive definitions for addition and multiplication, and (b) showing that such definitions are justified by proving the Recursion Theorem.

(**819**) **Remark:** Digression on Frege's Theorem. Frege's Theorem is that the Dedekind/Peano axioms for number theory are derivable in second-order logic extended solely by Hume's Principle. We've reconstructed this theorem in the following form: the Dedekind/Peano axioms for number theory are derivable in object theory extended solely with axiom (800), which asserts that a certain logically-defined relation (i.e., a relation defined without appeal to mathematical primitives) exists. Moreover, we've derived the Dedekind/Peano axioms in such a way that adapts Frege's methods to a modal context but which nevertheless yields a unique number series that can be used to count objects at any possible world.

The principal way in which we've departed from Frege's method is that in Frege's system, there are no indiscernible objects and so every object whatsoever could be counted. When one reconstructs Frege's theorem in standard (non-modal) second-order logic extended by Hume's Principle, the natural numbers he defines can count everything in the domain. Frege no doubt assumed that everything in the domain was discernible. But Frege didn't distinguish or axiomatize abstract objects, and in a general axiomatization that implies the existence of indiscernible abstracta, the entities that correspond to the Frege numbers count discernible objects. Thus, the fact that Frege's Theorem has been reconstructed relative to the discernible objects does not strike us as much of a limitation.

(**820**) **Remark:** A Potential But Unfounded Worry. In light of the theorems in (268), the new axiom asserted in (800) and definition of $\mathbb{P}$ stipulated in (801.1) have a consequence that raises some questions, though fortunately ones that can be put to rest. Recall that (268.1) and (268.2) assert:

$$\forall G \exists x \exists y (A!x \,\&\, A!y \,\&\, x \neq y \,\&\, [\lambda z\, Gzx] = [\lambda z\, Gzy]) \tag{268.1}$$

$$\forall G \exists x \exists y (A!x \,\&\, A!y \,\&\, x \neq y \,\&\, [\lambda z\, Gxz] = [\lambda z\, Gyz]) \tag{268.2}$$

These imply, respectively, that:

$$\exists x \exists y (A!x \,\&\, A!y \,\&\, x \neq y \,\&\, [\lambda z\, \mathbb{P}zx] = [\lambda z\, \mathbb{P}zy])$$

$$\exists x \exists y (A!x \,\&\, A!y \,\&\, x \neq y \,\&\, [\lambda z\, \mathbb{P}xz] = [\lambda z\, \mathbb{P}yz])$$

Let a and b be witnesses to the first existential claim, and c and d be witnesses to the second. That is suppose:

(ϑ) $A!a \,\&\, A!b \,\&\, a \neq b \,\&\, [\lambda z\, \mathbb{P}za] = [\lambda z\, \mathbb{P}zb]$

(ζ) $A!c \,\&\, A!d \,\&\, c \neq d \,\&\, [\lambda z\, \mathbb{P}cz] = [\lambda z\, \mathbb{P}dz]$

These consequences give rise to the following questions: (a) is anything a predecessor of either a or b, and (b) are c or d predecessors of anything? If the answer to the first is yes, then we could derive a contradiction from (ϑ) and the fact $\mathbb{P}$ is functional, i.e., $\forall x \forall y \forall z (\mathbb{P}xy \,\&\, \mathbb{P}xz \rightarrow y = z)$ (802.4):

> Suppose $\exists y \mathbb{P}ya$. Let e be such an object, so that $\mathbb{P}ea$. Then by β-Conversion, $[\lambda z\, \mathbb{P}za]e$. So by ($\vartheta$), $[\lambda z\, \mathbb{P}zb]e$. By β-Conversion, $\mathbb{P}eb$. But since (ϑ) also implies $a \neq b$, we now we have $\mathbb{P}ea$, $\mathbb{P}eb$ and $a \neq b$. This contradicts the functionality of $\mathbb{P}$.

Similarly, if c or d are predecessors of anything, then then we could derive a contradiction from (ζ) and the fact (802.3) that $\mathbb{P}$ is one-to-one (exercise).

But these derivations are simply *reductio* arguments to the conclusions that nothing is an immediate predecessor of a and b (i.e., that $[\lambda z\, \mathbb{P}za]$ and $[\lambda z\, \mathbb{P}zb]$ are unexemplified), and that c and d aren't immediate predecessors of anything (i.e., that $[\lambda z\, \mathbb{P}cz]$ and $[\lambda z\, \mathbb{P}dz]$ are unexemplified). More generally, we can conclude that none of a, b, c, and d can be natural numbers.

14.7 Number Theory

(**821**) **Definition:** Notation for Successors. By theorem (818) we know that every natural number has a unique successor among the natural numbers. Hence, we may introduce the notation n' to abbreviate the definite description *the natural number of which n is a predecessor*:

$$n' =_{df} \imath m \mathbb{P}nm$$

We henceforth refer to n' as *the successor of* n.

> **Note:** Since we are now using the prime symbol as a term-forming operator on terms that denote natural numbers, we must henceforth *refrain* from using prime symbols on our restricted variables n, m, and k to form new variables. However, we may continue to use prime notation on other variables — x and x' may be used as distinct general variables, and u and u' as distinct variables for discernible objects.

(**822**) **Theorem:** Fact About Successors. If n is identical to m, then the successor of n is identical to the successor of m:

$$n = m \rightarrow n' = m'$$

(**823**) **Theorem:** Fact Underlying Induction Using Successor Notation. It is common mathematical practice to inductively define a sequence of numerically-indexed notions. In these definitions, two clauses are used: in the first clause, the notion is defined for the index Zero, and in the second clause, the notion is defined for the index n' in terms of the notion for index n. Though the Principle of Mathematical Induction ultimately grounds this practice, the use of successor notation often simplifies the formulation of the notions being defined inductively. The theorem that grounds this practice of using successor notation can now be stated very simply, namely, a natural number is either identical to Zero or identical to n', for some natural number n:

$$m = 0 \vee \exists n(m = n')$$

Study of the proof of this theorem shows that in addition to the definitions of $\mathbb{N}$, $\mathbb{P}$, $\mathbb{P}^*$, $\mathbb{P}^+$, $=_D$, and $'$, the result depends on a fact about the weak ancestral of *predecessor* (795.7), and the one-to-one and functional character of *predecessor*, (802.3) and (802.4), respectively.

(**824**) **Theorem:** Natural Numbers Are Predecessors of Their Successors. By (153.2), it is a consequence of definition (821) that a natural number n is a predecessor of its successor:

$$\mathbb{P}nn'$$

(**825**) **Definitions:** Introduction of the Numerals. Since every natural number has a unique successor, we may introduce the (base 10) numerals '1', '2', '3', ... , as abbreviations, respectively, for the descriptions *the successor of 0*, *the successor of 1*, *the successor of 2*, etc.

(.1) $1 =_{df} 0'$

(.2) $2 =_{df} 1'$

(.3) $3 =_{df} 2'$

$\vdots$

The ellipsis is to be continued by a sequence of definitions with analogous definienda and definientia, ordered according to the base 10 representation of the natural numbers. Note that the definientia of the terms being introduced here are all significant. In English, the new definienda may be read, respectively: One, Two, Three, etc.

(**826**) **Theorems:** Numerical Order. The *predecessor* ordering of the natural numbers now follows directly from (825) and (824):

(.1) $\mathbb{P}01$

(.2) $\mathbb{P}12$

(.3) $\mathbb{P}23$

$\vdots$

(**827**) **Theorems:** Number-Theoretic Facts. The objects defined in (825) are natural numbers: (.1) One is a natural number; (.2) Two is a natural number; (.3) Three is a natural number; ...

(.1) $\mathbb{N}1$

(.2) $\mathbb{N}2$

(.3) $\mathbb{N}3$

$\vdots$

(**828**) **Definition:** Restricting Relations. We define (.1) *the restriction of* binary relation G *to* property F, written $G_{\upharpoonright F}$, to be the property *being an F-object that bears G to y*:

(.1) $G_{\upharpoonright F} =_{df} [\lambda xy\, Fx \,\&\, Gxy]$

So, for example, $=_{D\upharpoonright E!}$ is the relation $[\lambda xy\, E!x \,\&\, x =_D y]$, i.e., the identity$_D$ relation restricted to the concrete objects.

More generally, (.2) *the restriction of* n'-ary relation G *to* n-ary relation S^n, written $G_{\upharpoonright S^n}$, is the n'-ary relation *being* objects $x_1,\ldots,x_n,y$ *such that* $x_1,\ldots,x_n$ *both exemplify* S^n *and bear* G *to* y:

(.2) $G_{\upharpoonright S^n} =_{df} [\lambda x_1 \ldots x_n y\, S^n x_1 \ldots x_n \,\&\, Gx_1 \ldots x_n y]$

In the case where $n = 0$, the definition yields that (.3) *the restriction of* property G *to* proposition p is the property *being such that* p *is true and* y *exemplifies* G:

(.3) $G_{\upharpoonright p} =_{df} [\lambda y\, p \,\&\, Gy]$

(**829**) **Theorem:** Fact About Restrictions of Rigid Relations. It is an interesting (and useful) fact that if an n'-ary relation G is rigid, then its restriction to a rigid n-ary relation S is rigid:

$$Rigid(G) \,\&\, Rigid(S^n) \rightarrow Rigid(G_{\upharpoonright S^n}) \qquad (n \geq 0)$$

(**830**) **Definitions:** Some Number-Theoretic Relations. We now want to introduce the relations *less than* and *less than or equal to*. We could define them, respectively, as $[\lambda xy\, \mathbb{N}x \,\&\, \mathbb{N}y \,\&\, \mathbb{P}^*xy]$ and $[\lambda xy\, \mathbb{N}x \,\&\, \mathbb{N}y \,\&\, \mathbb{P}^+xy]$. However, we can define them more simply as restricted binary relations (828.1) if we take advantage of the facts that $\mathbb{P}^*ny \rightarrow \mathbb{N}y$ (814.2) and $\mathbb{P}^+ny \rightarrow \mathbb{N}y$ (814.3). For then we may define (.1) *less than* is *the restriction of the ancestral of predecessor to* *being a natural number*; and (.2) *less than or equal to* is *the restriction of the weak ancestral of predecessor to being a natural number*. In both cases, the definiens is significant. We may therefore introduce the standard mathematical symbols for these number-theoretic relations:

(.1) $< =_{df} \mathbb{P}^*_{\restriction\mathbb{N}}$

(.2) $\leq =_{df} \mathbb{P}^+_{\restriction\mathbb{N}}$

Note that we restrict $<$ and $\leq$ to relate only natural numbers. Later, we will see that there are natural cardinals that are not natural numbers. Such cardinals could be related by $\mathbb{P}$ but will not fall under $<$ or $\leq$.

Given these definitions, we define (.3) *greater than* is the converse of *less than*, and (.4) *greater than or equal to* is the converse of *less than or equal to*. We introduce the standard mathematical symbols for these number-theoretic relations:

(.3) $> =_{df} [\lambda xy\, y < x]$

(.4) $\geq =_{df} [\lambda xy\, y \leq x]$

In the usual manner, we use infix notation for all of these new symbols.

(**831**) **Theorems:** Basic Facts About *Less Than* and *Less Than Or Equal To*. The following are easy consequences of the previous definitions and prior theorems. (.1) x is less than y if and only if x and y are both natural numbers and x is a predecessor-ancestor of y; and (.2) x is less than or equal to y if and only if x and y are both natural numbers and x is a weak predecessor-ancestor of y:

(.1) $x < y \equiv \mathbb{N}x \,\&\, \mathbb{N}y \,\&\, \mathbb{P}^*xy$

(.2) $x \leq y \equiv \mathbb{N}x \,\&\, \mathbb{N}y \,\&\, \mathbb{P}^+xy$

(**832**) **Theorems:** Numerical Facts About Inequalities. If we now focus on the natural numbers, then the preceding theorems immediately yield, for natural numbers n and m, that (.1) n is less than m iff n is a predecessor ancestor of m; (.2) n is less than or equal to m iff n is a weak predecessor ancestor of m; (.3) n is greater than m iff m is less than n; and (.4) n is greater than or equal to m iff m is less than or equal to n:

(.1) $n < m \equiv \mathbb{P}^*nm$

(.2) $n \leq m \equiv \mathbb{P}^+nm$

(.3) $n > m \equiv m < n$

(.4) $n \geq m \equiv m \leq n$

In light of these results, we restrict our attention below primarily to the basic theorems about $<$ and $\leq$.

(**833**) **Theorem:** Some Facts About Less Than (or Equal To). (.1) If n is a predecessor of m, then n is less than m; (.2) n is less than or equal to n; (.3) if n is less than m, then n is less than or equal to m; (.4) if n is less than or equal to m and n is not equal to m, then n is less than m; (.5) if n is less than m and m is less than k, then n is less than k; and (.6) if n is less than or equal to m and m is less than or equal to k, then n is less than or equal to k; (.7) if n is less than m and m is less than or equal to k, then n is less than k; (.8) if n is less than or equal to m and m is less than k, then n is less than k:

(.1) $\mathbb{P}nm \rightarrow n < m$

(.2) $n \leq n$

(.3) $n < m \rightarrow n \leq m$

(.4) $n \leq m \;\&\; n \neq m \rightarrow n < m$

(.5) $n < m \;\&\; m < k \rightarrow n < k$

(.6) $n \leq m \;\&\; m \leq k \rightarrow n \leq k$

(.7) $n < m \;\&\; m \leq k \rightarrow n < k$

(.8) $n \leq m \;\&\; m < k \rightarrow n < k$

(**834**) **Theorems:** Some Additional Facts About Less Than (or Equal To). (.1) n is less than n'; and (.2) n is less than or equal to n'; (.3) if One is less than or equal to n, then Zero is less than n; (.4) If n is less than m', then m is less than n; and (.5) n' is greater than Zero:

(.1) $n < n'$

(.2) $n \leq n'$

(.3) $1 \leq n \rightarrow 0 < n$

(.4) $m' \leq n \rightarrow m < n$

(.5) $n' > 0$

(**835**) **Theorem:** Numbering and Less Than or Equal To. Intuitively, if all discernible F things are G things, then the number of discernible F things should be less than or equal to the number of discernible G things. More precisely, if n numbers F, m numbers G, and every discernible F is a discernible G, then n is less than or equal to m:

$$\big(Numbers(n,F) \,\&\, Numbers(m,G) \,\&\, \forall u(Fu \to Gu)\big) \to n \le m$$

This fact is important when we prove that there is an infinite cardinal.

(**836**) **Theorems:** Numerical Facts About Less Than. Clearly, from (834.1) and the definitions of the numerals in (825), the following claims become modally strict theorems:

(.1) $0 < 1$

(.2) $1 < 2$

(.3) $2 < 3$

$\vdots$

The following may prove useful:

> **Exercise 6**. Note that if we can derive the following claims:
>
> $0' \le 1$, $0' \le 2$, ...
> $1' \le 2$, $1' \le 3$, ...
> $2' \le 3$, $2' \le 4$, ...
> $\vdots$
>
> then by (834.4), the following claims become modally strict theorems:
>
> $0 < 1$, $0 < 2$, ...
> $1 < 2$, $1 < 3$, ...
> $2 < 3$, $2 < 4$, ...
> $\vdots$
>
> So as an exercise, derive the first two of the former group, i.e., that $0' \le 1$ and $0' \le 2$.

(**837**) **Remark:** Necessity and the Number of Planets. In a system such as ours, where all the terms are rigid designators, it is easy to see why one of the Quinean arguments that cast suspicion on modal contexts can be easily undermined. Quine (1953 [1961], 143) noted that the following argument, stated in ordinary natural language, is invalid (soundness is not in question, since the second premise is no longer thought to be true):

> Necessarily, nine is greater than seven.
> Nine is (identical to) the number of planets.
> ∴ Necessarily, the number of planets is greater than seven.

Clearly, if the identity claim in the second premise has a rigid term ('nine') on one side of the identity claim and a non-rigid term ('the number of planets') on the other side, one can't expect to validly use the substitution of identicals to substitute one term for the other in the modal context provided by the first premise.

But if we represent the argument by using the rigidly designating definite descriptions of *applied* object theory, we have the following valid argument:

$\Box 9 > 7$
$9 = \imath x Numbers(x, P)$
$\therefore\ \Box \imath x Numbers(x, P) > 7.$

We can see why this is valid if we reason intuitively in the formal mode with primitive possible worlds: from the facts that (a) in every possible world, Nine is greater than Seven and (b) Nine is identical to the object that (at the actual world) numbers the planets, it follows that in every possible world, the object that (at the actual world) numbers the planets is greater than Seven.

So, as far as applied object theory goes, Quine's argument does nothing to undermine the logical rigor of modal contexts. Of course, if one wants to give a reading of the natural language argument that shows *why* it is invalid, one has several options: (1) eliminate, in the manner of Russell's theory, the description from the second premise and also eliminate it from the conclusion but with narrow scope with respect to the $\Box$,[375] or (2) introduce a new kind of definite description, namely, a non-rigid one and restrict the principle of the substitution of identicals so that identicals can't be substituted within modal contexts.[376]

[375]That is, one could suggest that the following is a reading of the argument on which it is invalid:

$\Box 9 > 7$
$\exists! x(Numbers(x, P)\ \&\ 9 = x)$
$\therefore\ \Box \exists! x(Numbers(x, P)\ \&\ x > 7)$

This is invalid, as can be seen if we reason intuitively in the formal mode with primitive possible worlds: from the facts that (a) in every possible world Nine is greater than Seven and (b) there is a unique object that, at the actual world, numbers the planets and is identical to Nine, it doesn't follow that in every possible world, there is a unique object that numbers the planets and is greater than Seven.

[376]That is, one could suggest that the following is a reading of the argument on which it is invalid, where we use the Greek ι instead of the symbol $\imath$ to form non-rigid descriptions of the form $\iota x \varphi$ (these are subject to restrictions on substitution, which are violated in the following argument):

$\Box 9 > 7$
$9 = \iota x Numbers(x, P)$
$\therefore\ \Box \iota x Numbers(x, P) > 7.$

We can see why this is invalid if we reason intuitively in the formal mode with primitive possible worlds: from the facts that (a) in every possible world, Nine is greater than Seven and (b) Nine is identical to the object that happens to number the planets, it doesn't follow that in every possible world $\boldsymbol{w}$, the object that numbers the planets at $\boldsymbol{w}$ is greater than Seven.

(**838**) **Theorems:** Further Facts About Natural Numbers. It is a consequence of our work thus far that (.1) a natural number is not less than itself; (.2) a natural number doesn't immediately precede itself; and (.3) a natural number is not identical to its own successor:

(.1) $n \not< n$

(.2) $\neg\mathbb{P}nn$

(.3) $n \neq n'$

(**839**) **Definition** and **Theorem:** Number-Identity. We now define a relation on the natural numbers and show it is a genuine, though restricted, notion of identity. To define the relation in question, recall the definition of the restriction of a relation G to property F in (828.1). Since $=_D$ is a relation, we may write its restriction to the property *being a natural number* ($\mathbb{N}$) as $=_{D\restriction\mathbb{N}}$. We therefore define *being number-identical*, written $\doteq$, as the relation *being identical*$_D$ restricted to the property *being a natural number*:

(.1) $\doteq \ =_{df} \ =_{D\restriction\mathbb{N}}$

We henceforth use $\doteq$ as infix notation. It follows immediately from this definition and the previous lemma that (.2) x is number-identical to y if and only if x and y are identical natural numbers:

(.2) $x \doteq y \equiv \mathbb{N}x \,\&\, \mathbb{N}y \,\&\, x = y$

We now show that $\doteq$ is a genuine, though restricted, notion of identity.

(**840**) **Theorem:** *Number-Identity* and Identity. It is now provable that (.1) if x is number-identical to y, then x is identical to y; (.2) if either x or y is a natural number, then x and y are number-identical if and only if they are identical; (.3) number-identity is reflexive; (.4) symmetric; (.5) transitive; and (.6) rigid; and (.7) n is less than or equal to m iff either n is less than m or n is identical to m:

(.1) $x \doteq y \rightarrow x = y$

(.2) $(\mathbb{N}x \vee \mathbb{N}y) \rightarrow (x \doteq y \equiv x = y)$

(.3) $n \doteq n$

(.4) $n \doteq m \rightarrow m \doteq n$

(.5) $(n \doteq m \,\&\, m \doteq k) \rightarrow n \doteq k$

(.6) $Rigid(\doteq)$

(.7) $n \leq m \equiv n < m \vee n = m$

So (.2) implies that natural numbers n and m are *number-identical* precisely when identical, and (.3) – (.5) tell us that *number-identity* is provably an equivalence relation on natural numbers. Notice that (.6) tells us that the necessity of identity applies to number-identity. (.7) shows that the traditional definition of *less than or equal to* is a theorem.

(**841**) **Definition** and **Theorem:** Positive Integers. We leave it as an exercise to show that the expression $[\lambda x\, x > 0)]$ is significant. Hence, we may define (.1) *being a positive integer*, written $\mathbb{N}_+$, as *being an x greater than Zero*:

(.1) $\mathbb{N}_+ =_{df} [\lambda x\, x > 0]$

It immediately follows that:

(.2) $\mathbb{N}_+ x \equiv x > 0$

(**842**) **Theorems:** One is a Positive Integer and Similar Facts. It now follows that (.1) One is a positive integer. More generally, (.2) if One is less than or equal to n, then n is a positive interger:

(.1) $\mathbb{N}_+ 1$

(.2) $1 \leq n \rightarrow \mathbb{N}_+ n$

(**843**) **Theorem:** Every Positive Integer Succeeds a Unique Natural Number. If x is a positive integer, there is a unique natural number that precedes it:

$$\mathbb{N}_+ x \rightarrow \exists! n \mathbb{P} n x$$

Cf. Frege 1884 (§78, Proposition 6) where he asserts that every Number except Zero follows in the series of natural numbers directly after a number.

(**844**) **Lemmas:** Facts About Numbering Successors. (.1) If discernible object u exemplifies F and n numbers the F-objects not identical$_D$ to u, then n' numbers F; and (.2) If n' numbers F, then for some discernible object u exemplifying F, n numbers the F-objects not identical$_D$ to u:

(.1) $Fu \,\&\, Numbers(n, F^{-u}) \rightarrow Numbers(n', F)$

(.2) $Numbers(n', F) \rightarrow \exists u(Fu \,\&\, Numbers(n, F^{-u}))$

(**845**) **Definition and Theorems:** The Exact Numerical Quantifiers for Discernible Objects. We can now define the *exact numerical quantifiers*, or *cardinality quantifiers*, for discernible objects in simple terms. For any number n, there are exactly n discernible objects that exemplify F, written $\exists!_n uFu$, if and only if n numbers F:

(.1) $\exists_{!n} uFu \ \equiv_{df}\ Numbers(n,F)$

Note that the definition of $\exists_{!n} uFu$ is a condition with the variables n and F free. Note also how the numerical subscript in (.1) compare with the numerical subscripts used to distinguish the variables and constants of the system. The latter are primitive metalinguistic expressions (as providing a convenient supply of distinct symbols); they are not numerals ranging over objects falling under the *concept* of number). But now we are justified in introducing numerical subscripts, on expressions and in definitions and theorems, as referencing numbers. Thus, (.1) makes use of our theoretically defined numbers. The quantifier is indexed by variable ranging over natural numbers, and that variable can be bound by a quantifier. For example, we can now state $\forall n \exists F \exists_{!n} uFu$. As we'll see, this is theorem (846.3). So the numerical expressions being used here do in fact range over the natural numbers that we've defined in the system. That is, we are now making use of *the natural numbers* to (exactly) count the discernible objects that fall under a property no matter what the modal context.

It follows from this definition that (.2) there are exactly 0 discernible objects exemplifying F if and only if there are no discernible objects exemplifying F, and (.3) there are exactly n' discernible objects exemplifying F if and only if there is a discernible object u such that u exemplifies F and such that there are n discernible objects that exemplify being an F-object other than u:

(.2) $\exists_{!0} uFu \equiv \neg \exists uFu)$

(.3) $\exists_{!n'} uFu \equiv \exists u(Fu \,\&\, \exists_{!n} vF^{-u}v)$

Thus, we've derived the classic inductive definition of the numerical quantifiers (cf. Mendelson 1964 [1997, p. 101, Exercise 2.71]). So, clearly, $\exists_{!n} uFu$ does in fact mean that there are exactly n distinct discernible F-exemplifiers.

(**846**) **Theorems:** Natural Numbers and Numerical Quantifiers Over Discernible Objects. It is also provable that (.1) n is the abstract object that encodes just the properties F such that there are exactly n discernible objects exemplifying F; and (.2) for every number n, there is a property F such that there are exactly n Fs:

(.1) $n = \iota x(A!x \,\&\, \forall F(xF \equiv \exists_{!n} uFu))$

(.2) $\forall n \exists F \exists_{!n} uFu$

Note how (.1) validates an idea put forward by Hodes in papers from 1984 and 1990. He summarizes his 1984 view, 'coding fictionalism', as the idea that "numbers are, loosely speaking, fictions created to encode cardinality-quantifiers, thereby clothing a certain higher-order logic in the attractive garments of lower-order logic" (1990, 350).[377] We validate this idea by taking his

[377] He goes on to say:

'fictions' to be 'reified logical patterns', i.e., abstract objects. Each equinumerosity$_D$ condition on properties in our logic could be represented in third-order logic as a property of properties. Thus, each natural number n encodes those properties F that fall under the following property of properties: being a property F such that there are exactly n discernible objects that exemplify F. One might note that Hodes's view holds, by contrast, for the numerical quantifiers over any objects, but this doesn't confer greater generality for his conception since we doubt his domains included indiscernible objects. In any case, he appeals to far richer assumptions to state his view, namely, the assumptions of third-order logic, model theory, applied set theory, etc. By contrast, object theory can express and derive a version of the view in a second-order logic that is interpretable under general Henkin models.

(**847**) **Remark:** General Numerical Quantifiers, or How to Count Abstracta. We can now make use of the natural numbers to inductively define exact numerical quantifiers for *any* objects, *any* relations, and for any condition φ. These quantifiers will be unlike the numerical quantifiers for discernible objects: we won't be able to transform assertions involving them into facts about the number indexing the quantifier. Nevertheless, the following definitions are precise and allow us to assert *there are exactly n entities* (individuals or relations, as the case may be) *such that* φ:

$$\exists_{!0}\alpha\varphi \ \equiv_{df} \ \neg\exists\alpha\varphi$$
$$\exists_{!n'}\alpha\varphi \ \equiv_{df} \ \exists\alpha(\varphi \,\&\, \exists_{!n}\beta(\varphi^{\beta}_{\alpha} \,\&\, \beta \neq \alpha))$$

> More precisely: arithmetic singular terms that appear to do the semantic job of designating numbers really do the different job of encoding cardinality-quantifiers; quantifier-phrases that appear to quantify over numbers really encode higher-order quantification over cardinality-quantifiers; predicate-phrases, whose logico-syntactic behavior make them of level one, really do the semantic work of expressions of higher levels.

His position was laid out earlier in 1984 where we find, on p. 143:

> In making what appears to be a statement about numbers one is really making a statement primarily about cardinality object-quantifiers; what appears to be a first-order theory about objects of a distinctive sort really is an encoding of a fragment of third-order logic.

And on p. 144:

> The mathematical-object picture may be described in two equivalent ways. ... or we may see it as a pretense of positing objects that intrinsically represent type 2 entities. This second description makes mathematical discourse, when carried on within the mathematical-object picture, a special sort of fictional discourse: numbers are fictions "created" with a special purpose, to encode numerical object-quantifiers and thereby enable us to "pull down" a fragment of third-order logic, dressing it in first-order clothing.

By 'higher-order logic', Hodes means third-order logic. In our system, the Comprehension Principle for Abstract Objects grounds the way we abstract the numbers from these higher-order facts.

As the simplest (non-vacuous) case of this definition, let α be the individual variable x and φ be Fx, so that we have:

$$\exists!_0 xFx \equiv_{df} \neg\exists xFx$$
$$\exists_{!n'} xFx \equiv_{df} \exists x(Fx \,\&\, \exists_{!n} y(Fy \,\&\, y \neq x))$$

Similarly, let α be the m-ary relation variable F^m (for some m, $m \geq 0$) and let φ be any formula. Then our definition yields:

$$\exists_{!0} F^m\varphi \equiv_{df} \neg\exists F^m\varphi$$
$$\exists_{!n'} F^m\varphi \equiv_{df} \exists F^m(\varphi \,\&\, \exists_{!n} G^m(\varphi^{G^m}_{F^m} \,\&\, G^m \neq F^m))$$

Though $\exists_{!n} xFx$ is not equivalent to $Numbers(n, F)$, it can be used to answer the question "How many abstract individuals are such that φ?" by showing, for some n, that $\exists_{!n} x(A!x \,\&\, \varphi)$, and we can answer the question "How many Fs are such that φ?" by showing, for some n, $\exists_{!n} F\varphi$. The answers make use of numerals as informative indices, and though these numerals can now be grounded in our theory of natural numbers, the answers to the questions just posed make no reference to the natural numbers once the defined terms expanded into primitive notation.

14.8 Functions and Recursive Definitions

It is now time to discuss functions in greater generality. Up to this point, we have defined what it is for a binary relation R to map F to G (743.1), and what it is for a relation R to map$_D$ F to G (749.1). But our goal in this section is to introduce and justify the usual recursive definitions of the arithmetic functions such as addition, multiplication, etc. To do this, we need to (a) define n-ary functions generally and understand how they behave, (b) define maps and mappings generally, i.e., say what it is for an n'-ary relation to be an n-ary mapping from relation S^n to property G and what it is for a relation to be a *restricted function* from S^n to G, (c) introduce and justify the definition of new relations by induction and new functions by recursion, and then (d) combine our investigations so that we can justifiably use recursive definitions to introduce the classical, recursive, number-theoretic functions used in Peano Arithmetic.

14.8.1 Total Functions

In what follows, we use $R, S, \ldots$ as variables ranging over relations. If we need arbitrary constants for relations, we'll numerically index the variables.

(**848**) **Definitions:** n-ary Total Functions. Let us say that (.1) binary relation R is a *unary total function*, written $Function^1(R)$, if and only if for every object x, there is a unique object y such that x and y exemplify R:

(.1) $Function^1(R) \equiv_{df} \forall x \exists ! y Rxy$

For convenience, we have dropped the superscript on R in $Function^1(R)$, since the arity of R can be inferred.

More generally, we say that (.2) an n'-ary relation R is an *n-ary total function*, written $Function^n(R)$, if and only if for any objects $x_1, \ldots, x_n$ there is a unique object y such that $x_1, \ldots, x_n, y$ exemplify R:

(.2) $Function^n(R) \equiv_{df} \forall x_1 \ldots \forall x_n \exists ! y R x_1 \ldots x_n y$ $\quad (n \geq 0)$

Again, we have dropped the superscript on R in $Function^n(R^{n'})$, since the arity of R must be n' when R is an n-ary total function. Thus, a *binary* total function is a tertiary relation R such that $\forall x \forall y \exists ! z Rxyz$. And so on.

Our definition applies even when $n = 0$ so that (.3) a *nullary* total function is any unary relation R that is uniquely exemplified:

(.3) $Function^0(R) \equiv_{df} \exists ! y Ry$

So any property exemplified by exactly one object is a nullary total function. Cf. Leinster (2014, 405), who stipulates that *elements are a special case of functions*.

(**849**) **Definitions:** A Distinguished Group of Necessary Relations. In the remainder of this chapter, we shall need to appeal to relations that are necessarily exemplified by every $x_1, \ldots, x_n$. Of course, we've already introduced a necessary property, namely L, which was defined in (203) as $[\lambda x\, E!x \rightarrow E!x]$ (*being concrete if concrete*). However, instead of generalizing this definition to relations of higher arity, it is simpler to introduce necessary n-ary relations for each arity n as follows. Recall that in (208), we let p_0 be the proposition $\forall x(E!x \rightarrow E!x)$. Then let us define the n-ary relation $\mathcal{U}^n$ to be the relation *being* $x_1, \ldots, x_n$ *such that* p_0:

$$\mathcal{U}^n =_{df} [\lambda x_1 \ldots x_n\, p_0]$$

When $n = 1$, this definition yields that $\mathcal{U}^1$ (henceforth $\mathcal{U}$) is the property $[\lambda x p_0]$, and when $n = 0$, the definition yields that $\mathcal{U}^0$ is the proposition $[\lambda\, p_0]$, which we know by (111.1) is just p_0.

(**850**) **Theorems:** Facts About $\mathcal{U}^n$. Clearly, it follows from β-Conversion and the fact that p_0 is a theorem that (.1) individuals $x_1, \ldots, x_n$ exemplify *being* $x_1, \ldots, x_n$ *such that* p_0. Hence (.2) individuals $x_1, \ldots, x_n$ exemplify $\mathcal{U}^n$. So by GEN, (.3) any individuals $x_1, \ldots, x_n$ exemplify $\mathcal{U}^n$, and, by RN, (.4) necessarily, any individuals $x_1, \ldots, x_n$ exemplify $\mathcal{U}^n$:

(.1) $[\lambda x_1 \ldots x_n\, p_0]x_1 \ldots x_n$

(.2) $\mathcal{U}^n x_1 \ldots x_n$

(.3) $\forall x_1 \ldots \forall x_n \mathcal{U}^n x_1 \ldots x_n$

(.4) $\Box \forall x_1 \ldots \forall x_n \mathcal{U}^n x_1 \ldots x_n$

So by (200.1), $\mathcal{U}^n$ is a necessary relation. When $n = 0$, (.1) asserts *that* p_0 *is true*; (.2) and (.3) assert $\mathcal{U}^0$ *is true*; and (.4) asserts *necessarily*, $\mathcal{U}^0$ *is true*. Finally, note that in these theorems, we may regard n as a free variable, so that, by GEN, each becomes a universally generalized theorem beginning *for every* n. Though we've previously regarded superscript numerals as mere decorations that indicate arity (given that the notion *natural number* is not a primitive of our language and isn't expressed in the primitive formulas of our language), we are now entitled, given that the notion *natural number* has been properly defined and introduced, to regard these superscript indices as more than mere decorations. Cf. the similar observation made about subscripts in the remarks just after definition (845.1).

(**851**) **Theorems:** Total Functions of Every Arity Exist. It is a theorem that: (.1) *being an x and y such that x exemplifies* $\mathcal{U}$ *and y is number-identical to Zero* is a unary total function; and (.2) *being an* $x_1, \ldots, x_n$ *and y such that* $x_1, \ldots, x_n$ *exemplify* $\mathcal{U}^n$ *and y is number-identical to Zero* is an n-ary total function:

(.1) $Function^1([\lambda xy\, \mathcal{U}x \;\&\; y \doteq 0])$

(.2) $Function^n([\lambda x_1 \ldots x_n y\, \mathcal{U}^n x_1 \ldots x_n \;\&\; y \doteq 0])$

Four observations may be of interest. First, in (.1), we may substitute any positive integer for Zero and the result remains a theorem. The same applies to (.2). Second, in (.2), we may regard n as a free variable, so that by GEN, it follows that the theorem holds for every n. Third, the theorem holds even when $n = 0$. In that case, the theorem asserts:

(.3) $Function^0([\lambda y\, y \doteq 0])$

Fourth, these theorems depend on axiom (800) and the definition of $\mathbb{P}$ (801.1). We can, however, prove that total functions of every arity exist without appealing to that axiom. See the next item.

(**852**) **Theorems:** Non-Numerical Total Functions of Every Arity Exist. Without relying on the axiom that *predecessor* is a relation, we can show that: (.1) some binary relations are unary total functions, and more generally, (.2) some n'-ary relations are n-ary total functions:

(.1) $\exists R(Function^1(R))$

(.2) $\exists R(Function^n(R))$

Again, in (.2), we may regard n as a free variable and so the theorem holds even when $n = 0$.

It might be instructive to discuss the proof strategy that establishes the existence of non-numerical total functions. The strategy starts from the fact that there are discernible objects ($\exists x D!x$), which follows *a fortiori* from (227.1), the T schema, and (273.4). Let a be such an object, i.e., assume $D!a$. Then, the binary relation, *being an x and y such that x exemplifies* $\mathcal{U}$ *and y is identical*$_D$ *to a*, i.e., $[\lambda xy\, \mathcal{U}x \,\&\, y =_D a]$, is a unary total function. Since this establishes that (.1) follows from $D!a$, it follows by $\exists$E that (.1) follows from $\exists x D!x$. But since the latter is a theorem, (.1) is a theorem. And by a similar argument, one could have reached this conclusion by considering the binary relation $[\lambda xy\, y =_D a]$.

Now for (.2), a similar proof strategy suffices, except in this case, the following n'-ary relation bears witness to (.2): *being an* $x_1,\ldots,x_n,y$ *such that* $x_1,\ldots,x_n$ exemplify $\mathcal{U}^n$ *and y is identical*$_D$ *to a*, i.e.,

$$[\lambda x_1 \ldots x_n y\, \mathcal{U}^n x_1 \ldots x_n \,\&\, y =_D a]$$

By reasoning similar to that described in the previous paragraph, (.2) is a theorem. And by a similar argument, the relation *being an* $x_1,\ldots,x_n,y$ *such that y is identical*$_D$ *to a*, i.e., $[\lambda x_1 \ldots x_n y\, y =_D a]$, could also have served as a witness and played a role in the proof of (.2).

(**853**) **Theorems:** Relations That are Necessarily Total Functions. It immediately follows by RN from (851.1) and (851.2), respectively, that (.1) it is necessarily the case that *being an x and y such that x exemplifies* $\mathcal{U}$ *and y is number-identical to Zero* is a unary total function; and (.2) it is necessarily the case that *being an* $x_1,\ldots,x_n$ *and y such that* $x_1,\ldots,x_n$ *exemplify* $\mathcal{U}$ *and y is number identical to Zero* is an n-ary total function:

(.1) $\Box Function^1([\lambda xy\, \mathcal{U}x \,\&\, y \doteq 0])$

(.2) $\Box Function^n([\lambda x_1 \ldots x_n y\, \mathcal{U}^n x_1 \ldots x_n \,\&\, y \doteq 0])$

Note that these theorems make an appeal to axiom (800), since $\doteq$ is defined in (839.1) as $=_{D\restriction\mathbb{N}}$, which in turn depends on a definition (801.1) of $\mathbb{P}$, the significance of which depends on axiom (800). But we need not have appealed to this axiom. The existence of non-numerical relations that are necessarily total functions is independent of this axiom.

To see this, recall the reasoning described above, after the statement of theorems (852.1) and (852.2). From the fact that there are discernible objects, we picked an arbitrary discernible object a and considered the relation $[\lambda xy\, \mathcal{U}x \,\&\, y =_D a]$. Analogous reasoning helps to establish the following fact, which is derivable without any number-theoretic axioms or theorems: (.3) some binary relations are necessarily unary total functions:

(.3) $\exists R \Box Function^1(R)$

Similarly, the relation we identified as a witness to (852.2) is also a witness to the fact that (.4) some n'-ary relations are necessarily n-ary total functions:

(.4) $\exists R \Box Function^n(R)$

(**854**) **Theorems:** Total Functions and Contingency. It is important to understand that our system implies: (.1) some relations are total functions but possibly fail to be total functions, and (.2) some relations fail to be total functions but possibly are total functions:

(.1) $\exists R(Function^n(R) \& \Diamond\neg Function^n(R))$

(.2) $\exists R(\neg Function^n(R) \& \Diamond Function^n(R))$

Clearly, the unary total function $[\lambda xy\, \mathcal{U}x \,\&\, y \doteq 0]$ that we discussed in (851.1) is not a witness to the $n = 1$ case of (.1) or (.2). As we saw, $[\lambda xy\, \mathcal{U}x \,\&\, y \doteq 0]$ is necessarily a total function, and so isn't a witness to (.1), and since it is in fact a total function, it isn't a witness to (.2).

Instead, the proof of (.1), for $n = 1$, starts from the fact that, by (217.1), there are contingently true propositions, i.e., that there are propositions p such that $p \,\&\, \Diamond\neg p$. Then we reason with respect to an arbitrary such contingent truth, say p_1, and show that the relation $[\lambda xy\, p_1 \,\&\, \mathcal{U}x \,\&\, y \doteq 0]$ is in fact a total function but possibly fails to be one. Call this relation f_1. The the argument that f_1 is a total function is a variant of the reasoning used in the proof of (852.1). The theorem is modally-strict because the hypothesis that p_1 is contingently true is discharged. Moreover, the argument that f_1 possibly fails to be a total function also starts from the fact that p_1 is contingently true.[378] So there are relations that are total functions only *contingently*.

Analogously, the proof of (.2) starts from the fact that there are contingently false propositions (217.2). But this proof is left as an exercise.

(**855**) **Remark:** Restricted Variables for Total Functions. To introduce restricted variables for total functions, we have to show, by (336), that $Function^n(R)$ is a *restriction condition*. Since it has a single free variable R, we only have to show that it is strictly non-empty and that it has strict existential import. To show that it is strictly non-empty, we have to show that $\exists R(Function^n(R))$ is a modally strict theorem, for any choice of n. But this follows from the theorems in (851). Finally, to show that it has strict existential import, we have to show that $Function^n(\Pi) \to \Pi\downarrow$, for every n-ary relation term Π, and for every choice

[378] Intuitively, since p_1 is contingently true, there is a possible world, say w, where p_1 is false. At w, no objects x and y exemplify f_1 since at w, no x and y are such that p_1. Hence, at w, it is not the case that for every x there is a unique y such that f_1xy. So f_1 is not a total function at w. Hence, it is possible that f_1 is not a total function.

of n. But this follows from the definitions in (848); for any n, if $Function^n(\Pi)$ holds, then Π is significant (otherwise the definiens fails to be true).

Although we've just established that $Function^n(R)$ is therefore a restriction condition, the theorems in (854) establish that $Function^n(R)$ is *not* a *rigid* restriction condition, as defined in (340). So though we may introduce restricted variables for total functions, we may not regard these variables as *rigid* restricted variables. So it is important to take care when reasoning with these variables in modal contexts, for the reasons mentioned in (340).

In what follows, then we use the following decorated, lower-case italic letters as restricted variables ranging over total functions:

$$\hat{f}, \hat{g}, \hat{h}, \ldots$$

The arity of the total function will be clear from the number of arguments. Consequently, expressions of the form $\hat{f}x_1 \ldots x_n y$ are simply n'-ary exemplification formulas, with the n-ary total function variable $\hat{f}$ being an n'-ary relation term. We turn next to some modal facts about total functions.

(**856**) **Definition:** Function Application. Given any total function $\hat{f}$, it follows by definition (848.2) that for any $x_1, \ldots, x_n$, there is a unique y such that $\hat{f}x_1 \ldots x_n y$. Hence the description $\imath y \hat{f}x_1 \ldots x_n y$ is significant. We may therefore define *the value of* the n-ary total function $\hat{f}$ *for* $x_1, \ldots, x_n$, written $\hat{f}(x_1, \ldots, x_n)$, to be *the* object y such that $\hat{f}x_1 \ldots x_n y$:

$$\hat{f}(x_1, \ldots, x_n) =_{df} \imath y \hat{f}x_1 \ldots x_n y$$

Thus, even $\hat{f}(\,)$ is well-defined: where $n = 0$, $\hat{f}$ is a nullary total function and $\hat{f}(\,)$ is defined as $\imath y \hat{f} y$.

In general, when $\hat{f}$ is an n-ary total function and $\kappa_1, \ldots, \kappa_n$ are any individual terms, the parentheses and any comma separators in the new term $\hat{f}(\kappa_1, \ldots, \kappa_n)$ jointly constitute an *individual term-forming operator*. Note that we can use our new terms to construct binary exemplification formulas such as $Ry\hat{f}(x_1, \ldots, x_n)$. Moreover, defined identity formulas such as $\hat{f}(x_1, \ldots, x_n) = y$ are also well-formed. Whenever such an identity formula is true and $n \geq 1$, we say that $\hat{f}$ *maps arguments* $x_1, \ldots, x_n$ *to the value* y. We therefore tolerate the ambiguity between saying $\hat{f}$ maps x to y and saying R maps F to G, as the latter was defined in (743.1), since the variables always make it clear which notion is being invoked. When $n = 0$ and $\hat{f}(\,) = y$, we simply say that the value of the nullary total function $\hat{f}$ is y.

Our definitions establish that the notion of function application is reducible to the notion of exemplification and the description operator *the*, the latter being governed by the classical theorems involving uniqueness (i.e., existence and identity) claims. Cf. Whitehead & Russell 1910–1913 [1925–1927], $*30 \cdot 01$ and $*30 \cdot 02$. Thus, our system stands in complete contrast to Frege's system of

1893/1903a, since there he took functions and function application as primitive, and identities of the form $\hat{f}(x_1,\ldots,x_n)=y$ as primitive atomic formulas.

(**857**) ⋆**Theorem:** Relations and Function Values. It is an immediate consequence of the preceding definition and the theory of descriptions that $\hat{f}$ maps the arguments $x_1,\ldots,x_n$ to the value y if and only if $x_1,\ldots,x_n,y$ exemplify $\hat{f}$:

$$\hat{f}(x_1,\ldots,x_n)=y \equiv \hat{f}x_1\ldots x_n y$$

It is important to understand why this theorem is not modally strict. Inspection of the proof makes it clear that ⋆-theorems (about definite descriptions) play a role in both directions of the biconditional. Intuitively, however, the situation is a familiar one. The theorem is a biconditional in which (a) the truth of the one condition (in this case, the right-hand condition) may vary from world to world, while (b) the truth of the other condition is rigidly tied to the identity of the object that in fact is the value of $\hat{f}$ for the arguments $x_1,\ldots,x_n$. So the equivalence is guaranteed to hold in fact but not necessarily.

(**858**) ⋆**Theorem:** Fact about Total Functions That Are Materially Equivalent Relations. It is a consequence of our definitions and theorems that if $\hat{f}$ and $\hat{h}$ are materially equivalent relations, then necessarily, for all objects $x_1,\ldots,x_n$, the value of $\hat{f}$ for $x_1,\ldots,x_n$ is identical to the value of $\hat{h}$ for $x_1,\ldots,x_n$:

$$\forall x_1\ldots\forall x_n\forall y(\hat{f}x_1\ldots x_n y \equiv \hat{h}x_1\ldots x_n y) \rightarrow \Box\forall x_1\ldots\forall x_n(\hat{f}(x_1,\ldots,x_n)=\hat{h}(x_1,\ldots,x_n))$$

Although the necessity of identity (125) plays a role in the proof, the reasoning nevertheless appeals to (857)⋆ and so fails to be modally strict.

(**859**) **Remark:** Total Functions Needn't Be Extensional Entities. In classical mathematics, functions $\hat{f}$ and $\hat{g}$ that map the same arguments to the same values are considered identical, i.e., the very same function. That is, classical mathematical functions obey the following principle of extensionality:

$$\forall x_1\ldots\forall x_n(\hat{f}(x_1,\ldots,x_n)=\hat{g}(x_1,\ldots,x_n)) \rightarrow \hat{f}=\hat{g}$$

This principle is not derivable in our system and we should *not* add it as an axiom, for it implies the identity of any relations that are materially equivalent total functions, i.e., it implies:

$$\forall x_1\ldots\forall x_n\forall y(\hat{f}x_1\ldots x_n y \equiv \hat{g}x_1\ldots x_n y) \rightarrow \hat{f}=\hat{g}$$

The proof is relatively straightforward.[379] Put aside for now the fact that the reasoning in the derivation appeals to (857)⋆ and so fails to be modally strict;

[379] Without loss of generality, we prove the unary case. Assume the principle of extensionality:

(E) $\forall x(\hat{f}(x)=\hat{g}(x)) \rightarrow \hat{f}=\hat{g}$

To show that this implies that $\hat{f}$ and $\hat{g}$ are identical when materially equivalent, assume that $\hat{f}$ and $\hat{g}$ are materially equivalent, i.e., further assume:

we'll discuss this in Remark (860). Since object theory (a) analyzes total functions as relations and (b) allows us to consistently assert the distinctness of materially equivalent relations if the case demands it, we may assert the distinctness of materially equivalent total functions when the need arises. But the principle of extensionality for total functions would undermine that ability and should therefore be excluded as a general principle. So total functions are not extensional entities in the classical sense of mathematical functions. Nevertheless, they do have well-defined identity conditions, namely, those given in (23.2) and (116.2). Interestingly, these identity conditions offer a *new* sense in which extensionally equivalent total functions are identical.[380]

(**860**) **Remark:** Observations. It is important to point out how the previous two theorems about total functions takes advantage of the definition of $\hat{f}(x)$ as the object y in fact such that $\hat{f}xy$. This definition actually helps one to *avoid errors*. It would be a mistake to assume (a) that total functions are necessarily total functions, or (b) that necessarily equivalent relations that happen to be total functions are total functions necessarily. To see (a), note that we wouldn't want to define $\hat{f}(x)$ non-rigidly for relations that are only contingently total functions, since then $\hat{f}(x)$ might fail to be defined when it appears in a modal context. So even if a relation $\hat{f}$ is a total function only contingently, $\hat{f}(x)$ is

---

$$\forall x \forall y(\hat{f}xy \equiv \hat{g}xy)$$

Now by two applications of ∀E, this second assumption implies:

(ϑ) $\hat{f}xy \equiv \hat{g}xy$

So we can show $\hat{f}(x)=y \equiv \hat{g}(x)=y$ as follows:

$$\begin{array}{rcll} \hat{f}(x)=y & \equiv & \hat{f}xy & \text{by (857)}\bigstar \\ & \equiv & \hat{g}xy & \text{by } (\vartheta) \\ & \equiv & \hat{g}(x)=y & \text{by (857)}\bigstar \end{array}$$

Since x and y aren't free in our second assumption, we may apply GEN twice to the last result to conclude:

(ζ) $\forall x \forall y(\hat{f}(x)=y \equiv \hat{g}(x)=y)$

Now, independently, by the commutativity of the biconditional, it follows from modally strict theorem (117.4) that $\forall y(x=y \equiv z=y) \equiv x=z$. Since this holds by GEN for any x and z, if we instantiate x in this result to $\hat{f}(x)$ and z to $\hat{g}(x)$, then we obtain as a modally strict fact:

$$\forall y(\hat{f}(x)=y \equiv \hat{g}(x)=y) \equiv \hat{f}(x)=\hat{g}(x)$$

Hence, by a Rule of Substitution, we may substitute $\hat{f}(x)=\hat{g}(x)$ for $\forall y(\hat{f}(x)=y \equiv \hat{g}(x)=y)$ in (ζ) to obtain $\forall x(\hat{f}(x)=\hat{g}(x))$. But, then, by the principle of extensionality assumed at the outset, $\hat{f}=\hat{g}$.

[380] To see this, consider unary total functions $\hat{f}$ and $\hat{g}$. Since they're both binary relations, (116.2) says that $\hat{f}$ and $\hat{g}$ are identical whenever, for any object x, $[\lambda y\, \hat{f}xy]=[\lambda y\, \hat{g}xy]$ and $[\lambda y\, \hat{f}yx]=[\lambda y\, \hat{g}yx]$. By (189), this requires both (a) that $[\lambda y\, \hat{f}xy]$ and $[\lambda y\, \hat{g}xy]$ are encoded by the same objects, and (b) that $[\lambda y\, \hat{f}yx]$ and $[\lambda y\, \hat{g}yx]$ are encoded by the same objects. Given a formal semantics in which the truth conditions of 'xF' are that the object denoted by 'x' is in the *encoding extension* of the property denoted by 'F', (116.1) guarantees that $\hat{f}$ and $\hat{g}$ are identical whenever the relational properties that $\hat{f}$ and $\hat{g}$ give rise to have the same encoding extension. So if we allow that having the same encoding extension is a way of being extensionally equivalent, then total functions $\hat{f}$ and $\hat{g}$ that are extensionally equivalent are identical.

defined no matter what the context. To see (b), recall the relation we labeled $\hat{f}_1$ that served as the witness to (854.1):

$$\hat{f}_1\text{: } [\lambda xy\, p_1 \,\&\, \mathcal{U}x \,\&\, y \doteq 0]$$

Compare $\hat{f}_1$ with:

$$\hat{h}_1\text{: } [\lambda xy\, p_1 \,\&\, y \doteq 0]$$

It is easy to show that $\Box\forall x\forall y(\hat{f}_1 xy \equiv \hat{h}_1 xy)$. But we shouldn't conclude that either $\hat{f}_1$ or $\hat{h}_1$ is a total function necessarily. So we should define $\hat{f}_1(x)$ and $\hat{h}_1(x)$ rigidly, since those values would be undefined in certain modal contexts. Intuitively, in worlds where $\hat{f}_1$ and $\hat{h}_1$ fail to be total functions, their values there are undefined.

Thus, the fact that important theorems governing function application fail to be modally strict is completely in line with expectations. Theorems (857)★, (858)★, and the reasoning in Remark (859) are precisely what we should expect given how function application has been defined.

14.8.2 Functions From Domains to Codomains

(**861**) **Remark:** Our Strategy. We now work our way towards a definition of the conditions under which a relation R is a function from F to G, and more generally, a function from S^n to G. In this Remark we motivate the elements of the definition by discussing only the simpler, less general case.

Intuitively, a relation R is a function from F to G when it satisfies three conditions:

1. R maps F to G, i.e., by (743.1), R relates each F-object to a unique G-object;
2. R relates F-objects only to G-objects; and
3. R doesn't relate any $\overline{F}$ objects to anything.

As we noted in (743.1), the first of the above conditions, i.e., R maps F to G (written $R\,|\!: F \longrightarrow G$), doesn't exclude either the possibility that R relates something (e.g., an F-object) to $\overline{G}$-objects or the possibility that R relates $\overline{F}$-objects to other objects. Indeed, as we saw in Figure 14.2, R maps F to G even if, additionally, it (a) relates F-objects to $\overline{G}$ objects, (b) relates $\overline{F}$-objects to G objects and (c) relates $\overline{F}$-objects to $\overline{G}$-objects.

But it is traditional to suppose that when R is a *function from F to G*, R doesn't relate anything to a non-G object and R doesn't relate a non-F thing to anything, since R has F as a domain and has G as a codomain. Consequently, in what follows, we work up to a definition of R is a function from F to G in stages. First, we generalize the notion R maps F to G to define the notion R maps relation S^n to G. Then, after working with this notion a bit, we'll say that R is *functional from* F to G just in case R maps F to G and R relates F-objects *only* to G-objects. (We'll note in passing that this implies that R is *functional on* F, i.e., implies that if an F-object x bears R to both y and z, then y is identical to z.) We'll then generalize the definition of *functional from* to a definition of R is functional from S^n to G. Finally, after working with these notions, we'll define R is a *function from F to G* just in case R is functional from F to G and R has F as a domain (and so doesn't relate $\overline{F}$-objects to anything). This definition will have captured the three conditions stated at the beginning of this Remark. Again, we'll generalize this to a definition of R is a function from S^n to G.

(**862**) **Definitions:** Maps From Relations to Properties. In (743.1), we defined R *maps* property F *to* property G:

$$R|\!: F \longrightarrow G \ \equiv_{df} \ \forall x(Fx \rightarrow \exists! y(Gy \,\&\, Rxy)) \qquad (743.1)$$

We now generalize this to relations that map n-ary *relations* to properties. Where R is an n'-ary relation, we say that R *maps* n-ary relation S *to* property G, written $R|\!: S^n \longrightarrow G$, just in case, whenever objects $x_1,\ldots,x_n$ exemplify S^n, there exists a unique y such that both y exemplifies G and $x_1,\ldots,x_n,y$ exemplify R:

(.1) $R|\!: S^n \longrightarrow G \ \equiv_{df} \ \forall x_1 \ldots \forall x_n(S^n x_1 \ldots x_n \rightarrow \exists! y(Gy \,\&\, Rx_1 \ldots x_n y))$

Where $n = 0$, this definition becomes (.2) R maps proposition p to property G just in case p implies that there exists a unique y such that y both exemplifies G and R:

(.2) $R|\!: p \longrightarrow G \equiv_{df} p \rightarrow \exists! y(Gy \,\&\, Ry)$

Thus, if p is a false proposition, then every relation R maps p to every property G, by failure of the antecedent. And if there is a unique object that is both G and R, then R maps every proposition (true or false) to G.

It is important to remember the caution posted in (743). If $R|\!: S^n \longrightarrow G$, for any n, we can't infer R is a function with domain S^n and codomain G. The above definition tells us nothing about whether R relates objects that exemplify $\overline{S^n}$ to anything, or whether R relates anything to $\overline{G}$-objects. *A fortiori*, we can't infer that R is a total function in the sense of (848.2).

(**863**) **Theorem:** Examples of Maps. It follows that (.1) the tertiary relation *being an x, y, and z such that x exemplifies $\mathcal{U}$ and y is identical$_D$ to z* maps the relation *being an x and y such that x exemplifies $\mathcal{U}$ and y is a discernible object* to the property *being discernible*; (.2) *number identity* maps *being a natural number* to *being a natural number*; and (.3) *being an x and y such that x exemplifies $\mathcal{U}$ and y is number-identical to Zero* maps the property $\mathcal{U}$ to the property $\mathbb{N}$:

(.1) $[\lambda xyz\, \mathcal{U}x \,\&\, y =_D z] |\!: [\lambda xy\, \mathcal{U}x \,\&\, D!y] \longrightarrow D!$

(.2) $\doteq |\!: \mathbb{N} \longrightarrow \mathbb{N}$

(.3) $[\lambda xy\, \mathcal{U}x \,\&\, y \doteq 0] |\!: \mathcal{U} \longrightarrow \mathbb{N}$

(**864**) **Theorem:** *Predecessor* is a Map. We also have the following distinguished case of a map, namely, *predecessor* maps *being a number* to *being a number*:

$\mathbb{P} |\!: \mathbb{N} \longrightarrow \mathbb{N}$

Some readers may find the following useful:

> **Exercise 9**. Show that $\neg(< |\!: \mathbb{N} \longrightarrow \mathbb{N})$ and $\neg(\leq |\!: \mathbb{N} \longrightarrow \mathbb{N})$, i.e., show that *less than* and *less than or equal to* do not map *being a number* to *being a number*. As part of the proof, use previous theorems to first show, for some n, m, and o, that $n < m$, $n < o$, and $m \neq o$ (and similarly for $\leq$).

(**865**) **Lemmas:** Fact About Restrictions on Relations. We note as a useful lemma that (.1) if R maps F to G, then the restriction of R to F maps F to G:

(.1) $R |\!: F \longrightarrow G \to R_{\restriction F} |\!: F \longrightarrow G$

More generally, (.2) if R maps S^n to G, then the restriction of R to S^n maps S^n to G:

(.2) $R |\!: S^n \longrightarrow G \to R_{\restriction S^n} |\!: S^n \longrightarrow G$

When $n = 0$, our theorem becomes (.3) if R maps p to G, then R restricted to p maps p to G:

(.3) $R |\!: p \longrightarrow G \to R_{\restriction p} |\!: p \longrightarrow G$

(**866**) **Definitions:** Functional From. Let R be a binary relation. Then we say: (.1) R is *functional from* F to G, written $R \dotdiv F \longrightarrow G$, if and only if R maps F to G and for any objects x and y, if x exemplifies F and R relates x to y, then y exemplifies G:

(.1) $R \dotdiv F \longrightarrow G \equiv_{df} R |\!: F \longrightarrow G \,\&\, \forall x \forall y (Fx \,\&\, Rxy \to Gy)$

Note that when R is functional from F to G, R does not relate F-objects to $\overline{G}$-objects, though R may still relate $\overline{F}$-objects to other things.

> **Example 9**. Suppose that there are exactly seven individuals $a - g$, and that the relevant facts about them are depicted in Figure 14.11: a and b are the only F-objects; c and d are the only G-objects; e exemplifies $\overline{F}$; f and g exemplify $\overline{G}$; and Rad, Rbd, Rec, and Ref. Then $R \dotdiv F \longrightarrow G$.

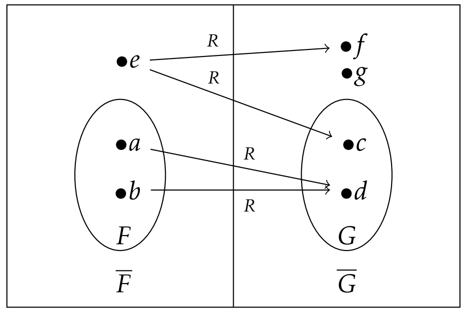


Figure 14.11: R is functional from F to G.

The following may be useful:

> **Exercise 10**. Compare Figure 14.11 with Figures 14.2 – 14.4, in which the Rs aren't functional from F to G. Say what minimal change could be made to all three earlier figures so that the Rs become functional from F to G.

With the example and exercises in mind, we say more generally (.2) R is *functional from S^n to G*, written $R \div S^n \longrightarrow G$, if and only if R maps S^n to G and for any objects $x_1,\ldots,x_n$ and y, if $x_1,\ldots,x_n$ exemplify S^n and R relates $x_1,\ldots,x_n$ to y, then y exemplifies G:

(.2) $R \div S^n \longrightarrow G \equiv_{df} R|\!:\!S^n \longrightarrow G \;\&\; \forall x_1 \ldots \forall x_n \forall y(S^n x_1 \ldots x_n \& R x_1 \ldots x_n y \rightarrow Gy)$

Note that when $n = 0$, (.2) asserts:

(.3) $R \div p \longrightarrow G \equiv_{df} R|\!:p \longrightarrow G \;\&\; \forall y(p \,\&\, Ry \rightarrow Gy)$

That is, R is functional from proposition p to property G whenever R maps p to G and for every y, if p is true and y exemplifies R, then y exemplifies G.

(**867**) **Theorems:** A Relation That Is Functional From $\mathbb{N}$ To $\mathbb{N}$. The following relation is functional from $\mathbb{N}$ to $\mathbb{N}$, namely, *being an x and y that are number-identical if natural numbers*:

$$[\lambda xy\, (\mathbb{N}x \,\&\, \mathbb{N}Ny) \rightarrow x \doteq y] \div \mathbb{N} \longrightarrow \mathbb{N}$$

Intuitively, this relation relates natural numbers to themselves and relates everything else to everything else. So although this relation is functional from $\mathbb{N}$ to $\mathbb{N}$, it is not yet a function from the natural numbers to the natural numbers given that objects which fail to be natural numbers are in the domain of R.

(**868**) **Theorems:** A Distinguished Relation That Is Functional From $\mathbb{N}$ To $\mathbb{N}$. It is also provable that *predecessor* is functional from $\mathbb{N}$ to $\mathbb{N}$:

$\mathbb{P} \mathrel{\dot{\div}} \mathbb{N} \longrightarrow \mathbb{N}$

(**869**) **Definitions:** Functional On. Let us say (.1) R is *functional on F*, written $FunctionalOn(R,F)$, just in case if an F-object bears R to any objects y and z, y and z are identical:

(.1) $FunctionalOn(R,F) \equiv_{df} \forall x(Fx \rightarrow \forall y \forall z(Rxy \,\&\, Rxz \rightarrow y = z))$

More generally, we say (.2) R is *functional on S^n*, written $FunctionalOn(R,S^n)$, just in case any objects $x_1,\ldots,x_n$ that exemplify S^n are such that any objects y and z that R relates $x_1,\ldots,x_n$ to are identical:

(.2) $FunctionalOn(R,S^n) \equiv_{df}$
$\forall x_1 \ldots \forall x_n(S^n x_1 \ldots x_n \rightarrow \forall y \forall z(Rx_1 \ldots x_n y \,\&\, Rx_1 \ldots x_n z \rightarrow y = z))$

Finally, when $n = 0$, our definition becomes (.3) R is *functional on p*, written $FunctionalOn(R,p)$, just in case p implies at most one thing exemplifies R:

(.3) $FunctionalOn(R,p) \equiv_{df} p \rightarrow \forall y \forall z(Ry \,\&\, Rz \rightarrow y = z))$

Note that every relation is functional on every false proposition.

(**870**) **Theorems:** Functional From Implies Functional On. It now follows that (.1) if R is functional from F to G, then R is functional on F:

(.1) $R \mathrel{\dot{\div}} F \longrightarrow G \rightarrow FunctionalOn(R,F)$

More generally, (.2) if R is functional from S^n to G, then R is functional on S^n:

(.2) $R \mathrel{\dot{\div}} S^n \longrightarrow G \rightarrow FunctionalOn(R,S^n)$

Finally, for $n = 0$, we have (.3) if R is functional from p to G, then R is functional on p:

(.3) $R \mathrel{\dot{\div}} p \longrightarrow G \rightarrow FunctionalOn(R,p)$

These results help to justify our use of the term *functional from* for the definienda in (866).

(**871**) **Remark:** Domains, Ranges, and Codomains. Our next goal is to define the conditions under which a relation R is a *function from F to G*, and more generally, when R is a function from S^n to G. Given the strategy outlined in Remark (861), R is a function from F to G whenever R is functional from F to G and R doesn't map any $\overline{F}$-objects to anything. This last condition can be captured by saying that R has F as a domain, where this means that F is exemplified by all and only the objects that R relates to something.

When we introduce the notion *R has F as a domain*, it is also natural to introduce the notion *R has H as a range*. Intuitively, R has H as a range whenever H

is exemplified by all and only the objects to which R relates something. Thus, these notions, of having a property as a domain or as a range, apply to every binary relation whatsoever — including those relations that aren't functional from F to G and that aren't maps from F to G. However, once we define R is a function from F to G by stipulating that R is functional from F to G and has F as a domain, we'll also stipulate that R has G as a codomain. The fact that R is a function from F to G doesn't require that every G-object has some F-object R related to it. So a codomain of R is not necessarily a range of R. Thus, our definitions will ensure that when R is a function from F to G, (a) R has F as a domain by definition, (b) R has G as codomain by definition, and (c) R has H as a range whenever H is exemplified by all and only the objects to which R relates something.

(**872**) **Definitions:** Domains and Ranges. Where R is a binary relation and F is a property, we say that (.1) *R has F as a domain* just in case F is exemplified by all and only the objects that bear R to something, and (.2) *R has H as a range* just in case H is exemplified by all and only the objects that something bears R to:

(.1) $HasDomain(R,F) \equiv_{df} \forall x(Fx \equiv \exists yRxy)$

(.2) $HasRange(R,H) \equiv_{df} \forall y(Hy \equiv \exists xRxy)$

As noted previously, these notions apply to any binary relation whatsoever — R need not be a total function nor a map from F to G.

More generally, where R is an n'-ary relation and S^n is an n-ary relation, we say that (.3) *R has S^n as a domain* just in case S^n is exemplified by all and only those objects $x_1,\dots,x_n$ that bear R to something and (.4) *R has H as a range* just in case H is exemplified by all and only those objects to which R relates some objects $x_1,\dots,x_n$:

(.3) $HasDomain(R,S^n) \equiv_{df} \forall x_1\dots\forall x_n(S^n x_1\dots x_n \equiv \exists yRx_1\dots x_n y)$

(.4) $HasRange(R,H) \equiv_{df} \forall y(Hy \equiv \exists x_1\dots\exists x_n Rx_1\dots x_n y)$

Analogously, these notions apply to any n'-ary R, even if R fails to be functional from S^n to G and fails to map S^n to G. Note also that when $n = 0$, our definitions assert (.5) R has p as a domain if and only if p is materially equivalent to the claim *something exemplifies* R, and (.6) R has H as a range just in case H is materially equivalent to R:

(.5) $HasDomain(R,p) \equiv_{df} p \equiv \exists yRy$

(.6) $HasRange(R,H) \equiv_{df} \forall y(Hy \equiv Ry)$

Notice here that if nothing exemplifies R, then R has every false proposition as a domain and has every unexemplified property as a range. These facts will become important later.

(**873**) **Remark:** Domains and Ranges Aren't Unique. Note that relations need not have unique domains or ranges. To see why, we need only consider, without loss of generality, a binary relation R that has Q, say, as a domain. (Analogous considerations apply both to n-ary relations generally and to any range that R has.) Then by (872.1), Q is exemplified by all and only those objects that bear R to something, i.e.,

$$\forall x(Qx \equiv \exists y Rxy)$$

Now consider the property $[\lambda x\, Qx \,\&\, p_0]$, where p_0 is $\forall x(E!x \rightarrow E!x)$. Call this property K. Though Q and K are materially, and indeed necessarily, equivalent, our theory doesn't require them to be identical. But it is straightforward to show that R also has K as a domain.[381] So if Q and K are distinct, R fails to have a unique domain.

> **Exercise 11**. Show (a) that R need not have a unique range, and (b) that if R is an empty binary relation, i.e., $\neg\exists x\exists y Rxy$, then R has F as a domain if and only if $\neg\exists x Fx$.

(**874**) **Definitions:** Functions From Domains to Codomains. We now say that (.1) R is a *function from* property F *to* property G if and only if R is functional from F to G and R has F as a domain:

(.1) $R : F \longrightarrow G \ \equiv_{df}\ R \dotdiv F \longrightarrow G \,\&\, HasDomain(R, F)$

When $R : F \longrightarrow G$, we say that R has G as a *codomain*. Thus, when R is a function from F to G, R doesn't relate $\overline{F}$-objects to anything, and doesn't relate anything to $\overline{G}$ objects.

> **Example 10**. Suppose that there are exactly 9 individuals a – i, and that the relevant facts about them are depicted in Figure 14.12: a, b, and c are the only F-objects; d, e, and f are the only G-objects; g exemplifies $\overline{F}$; h and i exemplify $\overline{G}$; and Rae, Rbe, and Rcf. Then $R : F \longrightarrow G$.

The following may prove useful:

[381] To see that R has K as a domain in this scenario, we start with the assumption that R has Q as a domain, i.e.,

(ϑ) $\forall x(Qx \equiv \exists y Rxy)$

Note independently that by β-Conversion and definition of K, it is a modally strict fact that $Kx \equiv (Qx \,\&\, p_0)$. And by propositional logic, it is a modally strict fact that $(Qx \,\&\, p_0) \equiv Qx$. Hence by modally strict reasoning, $Kx \equiv Qx$, which by the commutativity of the biconditional, yields $Qx \equiv Kx$. So by a Rule of Substitution, it follows from (ϑ) that $\forall x(Kx \equiv \exists y Rxy)$. But then by definition, R has K as a domain.

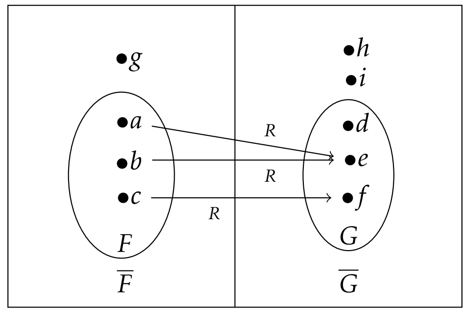


Figure 14.12: R is a function from F to G.

Exercise 12. Say exactly why R in Figure 14.12 would fail to be a function from F to G if either (a) R were to relate g to i, (b) R were to relate a to h, or (c) R were to relate a to d in addition to relating a to e.[382]

With this in mind, we may generalize (.1) by saying (.2) an n'-ary relation R is a *function from* relation S^n *to* property G, written $R : S^n \longrightarrow G$, just in case R is functional from S^n to G and R has S^n as a domain:

(.2) $R : S^n \longrightarrow G \ \equiv_{df} \ R \dotdiv S^n \longrightarrow G \ \& \ HasDomain(R, S^n)$

Similarly, when $R : S^n \longrightarrow G$, we say that R has G as a *codomain*. As with the case of binary functions, when R is a function from S^n to G, R doesn't relate $\overline{S^n}$-objects to anything, and doesn't relate anything to $\overline{G}$ objects.

When $n = 0$, our definition becomes (.3) R is a function from p to G just in case R is functional from p to G and R has p as a domain:

(.3) $R : p \longrightarrow G \ \equiv_{df} \ R \dotdiv p \longrightarrow G \ \& \ HasDomain(R, p)$

This definition yields an interesting consequence of the relational analysis of functions, namely, that there exist nullary functions from p to G that have every false proposition as a domain and every unexemplified property as a range. We'll prove this below, though some readers may wish to consider the matter for themselves beforehand. This consequence requires us to take precautions when we extend the definition of functional application to functions R from p to G, since when R is a nullary function with an empty range, it is a 'valueless' function.

[382]Solution: In case (a), R would fail to have F as a domain, and so would fail to be a function from F to G. In case (b), R would fail to relate F-objects only to G-objects, and hence would fail to be a function from F to G because it would fail the second clause of the definition of F is functional from F to G. In case (c), R would fail to be a function because it would fail to be a map from F to G — R would relate a to distinct G-objects d and e.

(**875**) **Theorems:** Summary Facts About Functions From S^n to G. Recall the discussion in Remark (861), where we outlined the three conditions under which R is a function from F to G. These conditions are now captured by the following summarizing facts: (.1) R is a unary function from F to G just in case (a) R relates each F-object to a unique G-object, (b) R relates F-objects only to G-objects, and (c) R relates x to something if and only if x exemplifies F; and more generally, (.2) R is an n-ary function from S^n to G just in case (a) R relates any n objects exemplifying S^n to a unique G-object, (b) R relates objects exemplifying S^n only to G-objects, and (c) R relates n objects to something if and only if those objects exemplify S^n:

(.1) $R : F \longrightarrow G \equiv$
$$\forall x(Fx \rightarrow \exists! y(Gy \,\&\, Rxy)) \,\&$$
$$\forall x \forall y(Fx \,\&\, Rxy \rightarrow Gy) \,\&$$
$$\forall x(Fx \equiv \exists y Rxy)$$

(.2) $R : S^n \longrightarrow G \equiv$
$$\forall x_1 \ldots \forall x_n(S^n x_1 \ldots x_n \rightarrow \exists! y(Gy \,\&\, Rx_1 \ldots x_n y)) \,\&$$
$$\forall x_1 \ldots \forall x_n \forall y(S^n x_1 \ldots x_n \,\&\, Rx_1 \ldots x_n y \rightarrow Gy) \,\&$$
$$\forall x_1 \ldots \forall x_n(S^n x_1 \ldots x_n \equiv \exists y Rx_1 \ldots x_n y)$$

When $n = 0$, our theorem becomes (.3) R is a nullary function from p to G just in case (a) p implies there is a unique object that exemplifies both G and R, (b) each object y such that p is true and Ry is such that Gy, and (c) p is true if and only if R is exemplified by something:

(.3) $R : p \longrightarrow G \equiv$
$$p \rightarrow \exists! y(Gy \,\&\, Ry) \,\&$$
$$\forall y(p \,\&\, Ry \rightarrow Gy) \,\&$$
$$p \equiv \exists y Ry$$

Moreover, we have (.4) if R is functional from F to G, then R restricted to F is a function from F to G; (.5) if R is functional from S^n to G, then R restricted to S^n is a function from S^n to G; and (.6) if R is functional from p to G, then R restricted to p is a function from p to G:

(.4) $R \dotdiv F \longrightarrow G \rightarrow R_{\restriction F} : F \longrightarrow G$

(.5) $R \dotdiv S^n \longrightarrow G \rightarrow R_{\restriction S^n} : S^n \longrightarrow G$

(.6) $R \dotdiv p \longrightarrow G \rightarrow R_{\restriction p} : p \longrightarrow G$

The following may prove useful to those of a mathematical bent, given how entrenched the notion of a *function on a domain* is in mathematics:

Exercise 13. Let us define: R is a *function on* (domain) F just in case $\exists G(R : F \longrightarrow G)$. Then, recall the definition of $FunctionalOn(R,F)$ (869) and prove that R is a function on F if and only if both $FunctionalOn(R,F)$ and $HasDomain(R,F)$. Show that this fact generalizes when we generalize the definition of *function on* to domains of arity $n \geq 2$ and arity $n = 0$.

Though the left-to-right direction of this exercise is straightforward, the right-to-left direction requires that one identify the witness to the existential claim. A proof is in the Appendix.

(**876**) **Theorems:** Examples of Functions From Domains to Codomains. It follows that (.1) the tertiary relation *being an x, y, and z such that x exemplifies $\mathcal{U}$ and y is identical$_D$ to z* is a function from the relation *being an x and y such that x exemplifies $\mathcal{U}$ and y is a discernible object* to the property *being discernible*; (.2) *number-identity* is a function from the property of *being a natural number* to *being a natural number*; (.3) the property *being a natural number that is number-identical to One* is a function from the proposition D^0 to the property *being a natural number that Zero immediately precedes*:

(.1) $[\lambda xyz\, \mathcal{U}x \,\&\, y =_D z] : [\lambda xy\, \mathcal{U}x \,\&\, D!y] \longrightarrow D!$

(.2) $\doteq : \mathbb{N} \longrightarrow \mathbb{N}$

(.3) $[\lambda n\, n \doteq 1] : \mathcal{U}^0 \longrightarrow [\lambda n\, \mathbb{P}0n]$

(**877**) **Theorem:** Facts About Domains. (.1) If R is a function from F to G and $\forall x(Fx \equiv Hx)$, then R is a function from H to G:

(.1) $(R : F \longrightarrow G \,\&\, \forall x(Fx \equiv Hx)) \rightarrow R : H \longrightarrow G$

More generally, (.2) if R is a function from S^n to G and $\forall x_1 \ldots \forall x_n(S^n x_1 \ldots x_n \equiv T^n x_1 \ldots x_n)$, then R is a function from T^n to G:

(.2) $(R : S^n \longrightarrow G \,\&\, \forall x_1 \ldots \forall x_n(S^n x_1 \ldots x_n \equiv T^n x_1 \ldots x_n)) \rightarrow R : T^n \longrightarrow G$

When $n = 0$, our theorem becomes (.3) if R is a function from p to G and $p \equiv q$, then R is a function from q to G:

(.3) $(R : p \longrightarrow G \,\&\, p \equiv q) \rightarrow R : q \longrightarrow G$

(**878**) **Theorem:** Facts About Codomains. (.1) If R is a function from F to G (i.e., with domain F and codomain G) and G materially implies H, then R is a function from F to H (and so has H as codomain as well as G):

(.1) $(R : F \longrightarrow G \,\&\, \forall x(Gx \rightarrow Hx)) \rightarrow R : F \longrightarrow H$

Clearly, this theorem holds even when $\neg\forall x(Hx \rightarrow Gx)$. Consider the following:

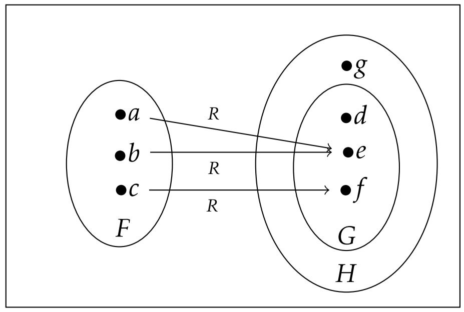


Figure 14.13: R is a function from F to G and from F to H, with distinct codomains G and H.

> **Example 11**. Suppose that there are exactly 7 individuals a – g, and that the relevant facts about them are depicted in Figure 14.13: a, b, and c are the only F-objects; d, e, and f are the only G-objects; d, e, f, and g are the only H-objects; and Rae, Rbe, and Rcf. Then R is a function from F to G, a function from F to H, and R has codomains G and H that aren't equivalent and, hence, not identical.

We noted earlier that when R is a function from F to G, R doesn't relate anything to $\overline{G}$ objects. So when (a) R is a function from F to G and (b) R is also a function from F to H because $\forall x(Gx \to Hx)$, then the only H-objects that R relates F-objects to are G-objects. This is also illustrated in Figure 14.13.

Our theorem generalizes for $n \geq 1$ and $n = 0$. (.2) If R is a function from S^n to G, and G materially implies H, then R is a function from S^n to H; and (.3) if R is a function from p to G, and G materially implies H, then R is a function from p to H:

(.2) $(R : S^n \longrightarrow G \;\&\; \forall x(Gx \to Hx)) \to R : S^n \longrightarrow H$

(.3) $(R : p \longrightarrow G \;\&\; \forall x(Gx \to Hx)) \to R : p \longrightarrow H$

(**879**) **Theorem** and **Definition:** A Distinguished Function with Distinct Codomain and Range. We now have as a theorem that *predecessor restricted to the natural numbers* is a unary function from $\mathbb{N}$ to $\mathbb{N}$ but has range *being a positive integer*:

(.1) $\mathbb{P}_{\restriction\mathbb{N}} : \mathbb{N} \longrightarrow \mathbb{N} \;\&\; \mathit{HasRange}(\mathbb{P}_{\restriction\mathbb{N}}, \mathbb{N}_+)$

So *predecessor restricted to the natural numbers* is a function that has $\mathbb{N}$ as a domain, $\mathbb{N}$ as a codomain, and $\mathbb{N}_+$ as a range. But what function is it? As a relation, we read $\mathbb{P}nm$ as *n is the predecessor of m*. But when we read $\mathbb{P}nm$ as a *function*, we have something that maps n to m, that is, it maps the predecessor of a number to the number that comes after. In other words, though this may be counterintuitive, we have here the *successor* function!

Though we already have a function symbol for successor, namely $'$, it will be useful in what follows to have a separate relation symbol that corresponds to the successor function. This will be especially helpful when we consider general composition of relations. Since it would be confusing (though accurate) to use the relation symbol $\mathbb{P}_{\restriction\mathbb{N}}$ for successor, we define:

(.2) $\boldsymbol{s} =_{df} \mathbb{P}_{\restriction\mathbb{N}}$

(**880**) **Theorems:** Facts About Total Functions. Recall both the definition of a total function in (848) and the definition of the property $\mathcal{U}^1$ (i.e., $\mathcal{U}$) in (849). It follows that (.1) R is a unary total function if and only if R is a function from $\mathcal{U}$ to $\mathcal{U}$:

(.1) $Function^1(R) \equiv R : \mathcal{U} \longrightarrow \mathcal{U}$

Now recall the definition of the relation $\mathcal{U}^n$ in (849. Then we have (.2) R is an n-ary total function if and only if R is a function from $\mathcal{U}^n$ to $\mathcal{U}$:

(.2) $Function^n(R) \equiv R : \mathcal{U}^n \longrightarrow \mathcal{U}$

Finally if we recall the definition of $\mathcal{U}^0$ and let $n = 0$, then (.2) asserts that (.3) R is a nullary total function if and only if R is a function from $\mathcal{U}^0$ to $\mathcal{U}$:

(.3) $Function^0(R) \equiv R : \mathcal{U}^0 \longrightarrow \mathcal{U}$

Thus, for every n, each n-ary total function R has relation $\mathcal{U}^n$ as domain and property $\mathcal{U}$ as codomain, and we can always regard R as a function from $\mathcal{U}^n$ to $\mathcal{U}$. Note also that (.3) has the interesting consequence that (.4) if R is a nullary total function, then for any proposition p, R has p as a domain if and only if p is true:

(.4) $Function^0(R) \rightarrow \forall p(HasDomain(R, p) \equiv p)$

(**881**) **Definitions:** One-to-One and Onto Functions From S^n to G. Recall that in (743.2) and (743.3), we defined, respectively, R maps F to G one-to-one and R maps F onto G. We could generalize those definitions to define R maps S^n to G one-to-one and R maps S^n onto G. And, similarly, we could generalize the definitions in (866) to define R is functional from S^n to G one-to-one, and R is functional from S^n onto G. But since our primary interest in what follows is in relations that are functions, and not just maps, we proceed directly to define one-to-one and onto functions from S^n to G.

Thus, we say (.1) R is a *one-to-one* function from F to G whenever R is a function from F to G and for any objects x, y, and z if x bears R to z and y bears R to z, then $x = y$:

(.1) $R : F \xrightarrow{\text{1-1}} G \equiv_{df} R : F \longrightarrow G \;\&\; \forall x \forall y \forall z (Rxz \;\&\; Ryz \rightarrow x = y)$

Compare this with the definition of $R|: F \xrightarrow{\text{1-1}} G$ in (743.2). When we know R is a function from F to G, we can simplify the second conjunct of the definiens. We don't need to require $\forall x \forall y \forall z((Fx \& Fy \& Gz) \rightarrow (Rxz \& Ryz) \rightarrow x = y)$ because the antecedent $Fx \;\&\; Fy \;\&\; Gz$ is implied when Rxz and Ryz. The definition of $R : F \longrightarrow G$ requires that R have F as a domain, and so it follows *a fortiori* from the definition of *HasDomain* in (872.1) that Rxz and Ryz imply, respectively, that Fx and Fy. And since R has to be functional from F to G to be a function from F to G, Fx and Rxz imply Gz.

More generally, (.2) R is a *one-to-one* function from S^n to G whenever R is a function from S^n to G and for any objects $x_1, \ldots, x_n, y_1, \ldots, y_n, z$, if $Rx_1 \ldots x_n z$ and $Ry_1 \ldots y_n z$, then $x_1 = y_1$ and ... and $x_n = y_n$:

(.2) $R : S^n \xrightarrow{\text{1-1}} G \equiv_{df} R : S^n \longrightarrow G \;\&$
$\forall x_1 \ldots \forall x_n \forall y_1 \ldots \forall y_n \forall z (Rx_1 \ldots x_n z \;\&\; Ry_1 \ldots y_n z \rightarrow (x_1 = y_1 \;\&\; \ldots \;\&\; x_n = y_n))$

When $n = 0$, (.2) reduces to:

(.3) $R : p \xrightarrow{\text{1-1}} G \equiv_{df} R : p \longrightarrow G \;\&\; \forall y \forall z (Ry \;\&\; Rz \rightarrow y = z)$

Now let us say (.4) R is a function from F *onto* G just in case R is a function from F to G and every G-object is an object to which R relates something:

(.4) $R : F \underset{\text{onto}}{\longrightarrow} G \equiv_{df} R : F \longrightarrow G \;\&\; \forall y (Gy \rightarrow \exists x Rxy)$

Compare this with the definition of $R|: F \underset{\text{onto}}{\longrightarrow} G$ in (743.3). When we know R is a function from F to G, we can simplify the second conjunct of the definiens. We don't need to require $\forall y (Gy \rightarrow \exists x (Fx \;\&\; Rxy))$ because Fx is implied when Rxy. Since we know by the definition of $R : F \longrightarrow G$ that R has F as a domain, it follows *a fortiori* from the definition of *HasDomain* in (872.1) that Rxy implies Fx.

More generally, (.5) R is a function from S^n *onto* G just in case R is a function from S^n to G and every G-object is an object to which R relates some things $x_1, \ldots x_n$:

(.5) $R : S^n \underset{\text{onto}}{\longrightarrow} G \equiv_{df} R : S^n \longrightarrow G \;\&\; \forall y (Gy \rightarrow \exists x_1 \ldots \exists x_n Rx_1 \ldots x_n y)$

When $n = 0$, we have:

(.6) $R : p \underset{\text{onto}}{\longrightarrow} G \equiv_{df} R : p \longrightarrow G \;\&\; \forall y (Gy \rightarrow Ry)$

The following may prove useful:

Exercise 14. Construct some specific examples of (.1) – (.6) above.

Note, finally, that since R is an n-ary *total* function iff R is a function from $\mathcal{U}^n$ to $\mathcal{U}$ (880.2), we may suppose that the case of n-ary one-to-one and onto total functions are covered by the above definitions. As a further exercise:

> **Exercise 15**. Prove that if R is a unary one-to-one function from $\mathcal{U}$ to $\mathcal{U}$ according to (.1) above, then R is a one-to-one relation according to definition (796.1), i.e., prove $R:\mathcal{U} \xrightarrow{\text{1-1}} \mathcal{U} \rightarrow \mathit{1\text{-}1}(R)$

(**882**) **Theorems:** Onto Functions From S^n to G and Ranges. It is a consequence of our definitions that (.1) a unary function from F onto G has G as a range; (.2) an n-ary function from S^n onto G has G as a range; and (.3) a nullary function from p onto G has G as a range:

(.1) $R:F \underset{\text{onto}}{\longrightarrow} G \rightarrow \mathit{HasRange}(R,G)$

(.2) $R:S^n \underset{\text{onto}}{\longrightarrow} G \rightarrow \mathit{HasRange}(R,G)$

(.3) $R:p \underset{\text{onto}}{\longrightarrow} G \rightarrow \mathit{HasRange}(R,G)$

(**883**) **Definition:** Reconstructing n^{th} Cartesian Products. Mathematicians often work with n-ary functions ($n \geq 1$) that have the n^{th} Cartesian product $(A_1 \times \cdots \times A_n)$ of the sets $A_1,\ldots,A_n$ as a domain and that have some set B as the range. Here, we reconstruct n^{th} Cartesian products as relations. For $n \geq 1$, we say that (.1) the n^{th} *Cartesian product* of properties $F_1,\ldots,F_n$, written $F_1 \times \cdots \times F_n$, is the relation *being objects* $x_1,\ldots,x_n$ *such that* x_1 *exemplifies* F_1 *and* ... *and* x_n *exemplifies* F_n:

(.1) $F_1 \times \cdots \times F_n \ =_{df}\ [\lambda x_1 \ldots x_n\, F_1x_1 \,\&\, \ldots \,\&\, F_nx_n]$

As an example, we have: $\mathcal{U} \times D! = [\lambda xy\, \mathcal{U}x \,\&\, D!y]$. Note that, given the above definition, $R:(F_1 \times \cdots \times F_n) \longrightarrow G$ becomes a special case of $R:S^n \longrightarrow G$. As a special case, for $n \geq 1$, we say (.2) the n^{th} *Cartesian product* of a property F, written $F^{\times n}$, is *being objects* $x_1,\ldots,x_n$ *such that* x_1 *exemplifies* F *and* ... *and* x_n *exemplifies* F:

(.2) $F^{\times n} =_{df}\ [\lambda x_1 \ldots x_n\, Fx_1 \,\&\, \ldots \,\&\, Fx_n]$ $(n \geq 1)$

It will also be helpful to define the edge or degenerate case, namely, the 0^{th} Cartesian product of a property F. Recall that we defined $\mathcal{U}^0$ in (849) as $[\lambda\, p_0]$, where p_0 is the necessary truth $\forall x(E!x \rightarrow E!x)$. So $\mathcal{U}^0$ is a true proposition. Hence we define (.3) the 0^{th} *Cartesian product* of a property F to be just $\mathcal{U}^0$:

(.3) $F^{\times 0} =_{df}\ \mathcal{U}^0$

To understand this definition, it may be helpful to point out that we shall want to use Cartesian products as domains for functions. So $F^{\times 0}$ needs to be the domain for a nullary function. Now if some relation R is a nullary function, then by (848.3), $\exists! y Ry$. Every such R has $\mathcal{U}^0$ (indeed, any true proposition) as a domain, by (880.4).

14.8.3 Restricted Functions and Functions Generally

(**884**) **Definitions:** Restricted Functions. Let us say R is a *unary restricted function*, written $\upharpoonright\text{-}\mathit{function}^1(R)$, just in case there is a property F and a property G such that R is a function from F to G but R is not a total function:

(.1) $\upharpoonright\text{-}\mathit{function}^1(R) \ \equiv_{df} \ \exists F \exists G(R : F \longrightarrow G \,\&\, \neg \mathit{Function}^1(R))$

In the usual manner, we have dropped the superscript on R since the arity of R can be inferred. More generally, we say (.2) R is an *n-ary restricted function*, written $\upharpoonright\text{-}\mathit{function}^n(R)$, just in case there is a relation S^n and a property G such that R is a function from S^n to G but R is not a total function:

(.2) $\upharpoonright\text{-}\mathit{function}^n(R) \ \equiv_{df} \ \exists S^n \exists G(R : S^n \longrightarrow G \,\&\, \neg \mathit{Function}^n(R))$

And when $n = 0$, our definition becomes (.3) R is an *nullary restricted function*, written $\upharpoonright\text{-}\mathit{function}^0(R)$, just in case there is a proposition p and a property G such that R is a function from p to G but R is not a total function:

(.3) $\upharpoonright\text{-}\mathit{function}^0(R) \ \equiv_{df} \ \exists p \exists G(R : p \longrightarrow G \,\&\, \neg \mathit{Function}^0(R))$

(**885**) **Theorems:** Facts About Restricted Functions. It now follows from previous definitions and theorems that (.1) R is a unary restricted function if and only if for some F and G, R is a function from F to G and F is not equivalent to the property $\mathcal{U}$:

(.1) $\upharpoonright\text{-}\mathit{function}^1(R) \equiv \exists F \exists G(R : F \longrightarrow G \,\&\, \neg \forall x(Fx \equiv \mathcal{U}x))$

More generally, (.2) R is an n-ary restricted function if and only if for some S^n and G, R is a function from S^n to G and S^n is not equivalent to the relation $\mathcal{U}^n$:

(.2) $\upharpoonright\text{-}\mathit{function}^n(R) \equiv \exists S^n \exists G(R : S^n \longrightarrow G \,\&\, \neg \forall x_1 \ldots \forall x_n(S^n x_1 \ldots x_n \equiv \mathcal{U}^n x_1 \ldots x_n))$

And when $n = 0$, our theorem becomes (.3) R is a nullary restricted function if and only if for some proposition p and property G, R is a function from p to G and p is not equivalent to $\mathcal{U}^0$:

(.3) $\upharpoonright\text{-}\mathit{function}^0(R) \equiv \exists p \exists G(R : p \longrightarrow G \,\&\, p \not\equiv \mathcal{U}^0)$

If we recall that by (849), (111.1), and a convention introduced in (208), $\mathcal{U}^0 = [\lambda p_0] = p_0 = \forall x(E!x \rightarrow E!x)$, then the above tells us that R is a nullary restricted function iff for some proposition p and property G, R is a function from p to G and p isn't equivalent to $\forall x(E!x \rightarrow E!x)$ (i.e., and p is false).

Finally, one other striking fact emerges, namely, (.4) R is a nullary restricted function if and only if R is an unexemplified property:

(.4) $\upharpoonright\text{-}function^0(R) \equiv \neg\exists y Ry$

(**886**) **Theorem:** Restricted Functions of Every Arity Exist. It now follows that (.1) unary restricted functions exist:

(.1) $\exists R(\upharpoonright\text{-}function^1(R))$

More generally, (.2) restricted functions of arity $n \geq 1$ exist:

(.2) $\exists R(\upharpoonright\text{-}function^n(R))$

When $n = 0$, our theorem becomes (.3) nullary restricted functions exist:

(.3) $\exists R(\upharpoonright\text{-}function^0(R))$

(**887**) **Theorems:** Specific Restricted Functions. (.1) *being an* x, y, *and* z *such that* y *is identical*$_D$ *to* z is a binary restricted function; (.2) *identity*$_D$ is a unary restricted function; (.3) *predecessor* is a unary restricted function; and (.4) *number-identity* is a unary restricted function:

(.1) $\upharpoonright\text{-}function^2([\lambda xyz\, y =_D z])$

(.2) $\upharpoonright\text{-}function^1(=_D)$

(.3) $\upharpoonright\text{-}function^1(\mathbb{P})$

(.4) $\upharpoonright\text{-}function^1(\doteq)$

(**888**) **Theorem:** Facts About Nullary Restricted Functions. It is a consequence of our definitions and theorems that (.1) if R is a nullary restricted function, then R has a proposition as a domain iff that proposition is false, and (.2) if R is a nullary restricted function, then R a property as a range iff that property is unexemplified:

(.1) $\upharpoonright\text{-}function^0(R) \rightarrow \forall p(HasDomain(R, p) \equiv \neg p)$

(.2) $\upharpoonright\text{-}function^0(R) \rightarrow \forall H(HasRange(R, H) \equiv \neg\exists x Hx)$

Contrast (.1) with an earlier fact (880.4) about total functions: nullary total functions have all and only true propositions as domains, whereas nullary restricted functions have all and only false propositions as domains.

(**889**) **Definitions:** Functions in Complete Generality. In what follows, we shall say, for $n \geq 0$, that a relation R is an n-ary *function*, written $function^n(R)$, just in case there is a relation S^n and a property G such that R is a function from S^n to G:

$$function^n(R) \equiv_{df} \exists S^n \exists G(R : S^n \longrightarrow G) \qquad (n \geq 0)$$

(**890**) **Theorem:** Facts About Functions Generally. R is a function if and only if R is a total function or R is a restricted function:

$$function^n(R) \equiv Function^n(R) \vee \upharpoonright\text{-}function^n(R)$$

Thus, when we talk about functions generally, we are talking about both total functions and restricted functions. As an exercise to test understanding, the following may prove useful:

> **Exercise 16**. Show that all nullary functions are one-to-one, i.e., that $function^0(R) \rightarrow \forall y \forall z(Ry \,\&\, Rz \rightarrow y = z)$.

(**891**) **Remark:** Restricted Variables for Functions Generally. We leave it as an exercise to show that $function^n(R)$ is a restriction condition on R but not a rigid restriction condition, as these notions were defined in (336) and (340). We therefore use the variables $\overline{f}, \overline{g}, \overline{h}, \ldots$ to range over functions generally, of whatever arity is needed.

(**892**) **Theorem:** Another Fact About Functions Generally. If there is an object y such that function $\overline{f}$ relates $x_1, \ldots, x_n$ to y, there is a unique object y such that $\overline{f}$ relates $x_1, \ldots, x_n$ to y:

$$\exists y \overline{f} x_1 \ldots x_n y \rightarrow \exists! y \overline{f} x_1 \ldots x_n y$$

This holds for all n and all functions, total or restricted. It holds by failure of the antecedent for nullary restricted functions because they are unexemplified (885.4) and so the antecedent of the above theorem is false.

(**893**) **Definitions:** Extended Function Application. Given the previous theorem, we may *extend* the notion of *functional application* defined in (856) so that it applies to any function provided that it does in fact have a value. Where $\overline{f}$ is any n-ary function such that $\exists y \overline{f} x_1 \ldots x_n y$, we say *the value of* $\overline{f}$ *for* $x_1, \ldots, x_n$, written $\overline{f}(x_1, \ldots, x_n)$, is *the* object y such that $\overline{f} x_1 \ldots x_n y$:

$$\overline{f}(x_1, \ldots, x_n) =_{df} \imath y \overline{f} x_1 \ldots x_n y$$

This definition holds only when $\overline{f}$ in fact has a value for its arguments, since it takes advantage of the primitive Rule of Definition by Identity (73). This rule guarantees that if the definiens (when applied to any arguments) provably fails to be significant, then the definiendum (when applied to those arguments) provably fails to be significant.

As noted previously, when $\overline{f}$ is a nullary restricted function, $\neg\exists y\overline{f}y$ (that is, nothing exemplifies $\overline{f}$). Hence, $\neg\exists!y\overline{f}y$ and so both $\imath x\overline{f}x$ and $\overline{f}(\,)$ are not significant. However, we also noted earlier in (856) that $\hat{f}(\,)$ is well-defined when $\hat{f}$ is a nullary total function. When we focus on total functions, then, $\hat{f}(\,)$ is $\imath y\hat{f}y$.

In general, then when $\overline{f}$ is any n-ary function, and $\kappa_1,\ldots,\kappa_n$ are any individual terms ($n \geq 0$) exemplifying a relation that $\overline{f}$ has as a domain (872.3), the parentheses and any comma separators in the term $\overline{f}(\kappa_1,\ldots,\kappa_n)$ jointly constitute an *individual term-forming operator*. In the remainder of this chapter, we shall primarily be discussing *restricted numerical functions* $\overline{f}$ that have domain $\mathbb{N}^{\times n}$ and codomain $\mathbb{N}$.

14.8.4 Numerical Operations

(**894**) **Definition** and **Theorem:** n-ary Numerical Operations on the Natural Numbers. When $n \geq 0$, we say that (.1) n'-ary relation R is an n-ary *numerical operation*, written $Op^n(R)$, if and only if R is rigid and a function from $\mathbb{N}^{\times n}$ to $\mathbb{N}$:

(.1) $Op^n(R) \equiv_{df} Rigid(R) \,\&\, R : \mathbb{N}^{\times n} \longrightarrow \mathbb{N}$ $\qquad (n \geq 0)$

Observe that when $n = 0$, a nullary operation is defined as a rigid relation that maps $\mathbb{N}^{\times 0}$ $(= \mathcal{U}^0)$ to $\mathbb{N}$.

Clearly, if $Op^n(R)$, then if we detach the second conjunct of its definiens, apply $\exists$I twice and appeal to definition (889), then it follows that R is a function:

(.2) $Op^n(R) \rightarrow function^n(R)$

So in the case when $n = 0$, a 0-ary numerical operation is a nullary function (whose unique value is a natural number).

(**895**) **Theorems:** Rigidity of Numerical Operations. The result of applying a numerical operation should be completely and necessarily determined by the given operands. Out definitions preserve this, since it is a modally strict theorem that for any n'-ary relation R, if R is a n-ary numerical operation, then it is so necessarily:

$$\forall R(Op^n(R) \rightarrow \Box Op^n(R))$$

Since this theorem is modally strict, RN yields $\Box\forall R(Op^n(R) \rightarrow \Box Op^n(R))$.

(**896**) **Lemma:** The Successor Function is a Numerical Operation. Recall that in (879), we defined the successor function $\boldsymbol{s}$ as the predecessor relation restricted to the natural numbers $\mathbb{P}_{\restriction\mathbb{N}}$. We now prove that $\boldsymbol{s}$ is a unary numerical operation:

$$Op^1(\boldsymbol{s})$$

As part of the proof of this theorem, one has to prove the rigidity of $\boldsymbol{s}$, when considered as a relation. This follows from the facts that $\mathbb{P}$ is rigid (802.2), $\mathbb{N}$ is rigid (809.2), and the restriction of a rigid n'-ary relation to a rigid n-ary relation is rigid (829).

(**897**) **Definitions and Theorems:** Constant Relations. For any natural numbers $n, m \geq 0$, we define the constant n'-ary relation $\mathcal{C}^{n'}_m$ as *being an* $x_1, \ldots, x_n$ *and* y *such that* $x_1, \ldots, x_n$ *are natural numbers and* y *is number-identical to* m:

(.1) $\mathcal{C}^{n'}_m =_{df} [\lambda x_1 \ldots x_n y\ \mathbb{N}x_1 \ \& \ \ldots \ \& \ \mathbb{N}x_n \ \& \ y \doteq m]$ $\quad (n, m \geq 0)$

Note that $\mathcal{C}^1_m$ is simply $[\lambda y\ y \doteq m]$. It now follows that $\mathcal{C}^{n'}_m$ is an n-ary numerical operation:

(.2) $Op^n(\mathcal{C}^{n'}_m)$

Note that we have no particular need for a Zero constant function, as we have already defined the natural number Zero. But, if needed, we have 'Zero functions' of every arity $n \geq 0$, since $\mathcal{C}^1_0$, $\mathcal{C}^2_0$, etc., are all instances of (.1).

(**898**) **Definitions and Theorem:** Projection Relations. It is customary to introduce, for each $i \geq 1$, some i-ary projection functions onto the k^{th} argument. In the present case, we will define these as i'-ary *relations* ($i \geq 1$) indexed by k ($1 \leq k < i'$) that hold among i' numbers whenever the k^{th} argument matches the final argument. That is, these projection relations can be seen as selecting the k^{th} argument. Hence, there is no unary projection relation (i.e., nullary projection function), since there is no argument to project. So in the simplest case (.1), π^2_1 is simply the binary relation of number-identity ($\doteq$), which relates any two numbers n and m whenever its final (i.e, second) argument m matches its first argument n. In the next case, (.2) π^3_1 and π^3_2 are tertiary projection relations such that π^3_1 relates any three numbers n_1, n_2, and m whenever m equals n_1; and π^3_2 relates any three numbers n_1, n_2, and m whenever m equals n_2. If we proceed in this manner, we end up defining (.3) a group of i'-ary *projection* relations $\pi^{i'}_k$, whereby $\pi^{i'}_k$ relates a string of i' arguments whenever the final argument matches the k^{th} argument in the string, for $1 \leq k < i'$:

(.1) $\pi^2_1 =_{df}\ \doteq$

(.2) $\pi^3_1 =_{df} [\lambda n_1 n_2 m\ m \dot{=} n_1]$
$\pi^3_2 =_{df} [\lambda n_1 n_2 m\ m \dot{=} n_2]$

(.3) $\pi^{i'}_1 =_{df} [\lambda n_1 \ldots n_i m\ m \dot{=} n_1]$
$\vdots$
$\pi^{i'}_k =_{df} [\lambda n_1 \ldots n_i m\ m \dot{=} n_k] \quad (1 \leq k \leq i)$
$\vdots$
$\pi^{i'}_i =_{df} [\lambda n_1 \ldots n_i m\ m \dot{=} n_i]$

It is now provable that $\pi^{i'}_k$ is a numerical i-ary operation ($i \geq 2$):

(.4) $Op^i(\pi^{i'}_k) \quad (1 \leq k \leq i)$

(**899**) **Lemma:** Numerical Operations Relate Natural Numbers. (.1) If R is an n-ary numerical operation, then for any $x_1, \ldots, x_n$ and y, if R relates $x_1, \ldots, x_n$ to y, all of $x_1, \ldots, x_n, y$ are natural numbers; (.2) if R is an n-ary numerical operation, then for any objects $x_1, \ldots, x_n$, and y, R relates $x_1, \ldots, x_n$ to y if and only if y is identical to the object to which R relates $x_1, \ldots, x_n$; and (.3) if R is an n-ary numerical operation, then for any natural numbers $m_1, \ldots, m_n$, then R relates $m_1, \ldots, m_n$ to a unique natural number:

(.1) $Op^n(R) \rightarrow \forall x_1 \ldots \forall x_n \forall y(Rx_1 \ldots x_n y \rightarrow \mathbb{N}x_1 \;\&\; \ldots \;\&\; \mathbb{N}x_n \;\&\; \mathbb{N}y) \qquad (n \geq 0)$

(.2) $Op^n(R) \rightarrow \forall x_1 \ldots \forall x_n \forall y(Rx_1 \ldots x_n y \equiv y = \imath y Rx_1 \ldots x_n y) \qquad (n \geq 0)$

(.3) $Op^n(R) \rightarrow \forall m_1 \ldots \forall m_n \exists! k Rm_1 \ldots m_n k \qquad (n \geq 0)$

(**900**) **Definitions:** Composed Relations. Suppose G and H are both binary relations. Then we define (.1) *G composed with H*, written $G \circ H$ or, when delimiters are needed, $[G \circ H]$, as the binary relation *being an x and y such that for some z, x bears H to z and z bears G to y*:

(.1) $G \circ H =_{df} [\lambda xy\, \exists z(Hxz \;\&\; Gzy)]$

Now let G be an m'-ary relation ($m \geq 1$) and $H_1, \ldots, H_m$ be n'-ary relations ($n \geq 0$). Then we define (.2) *G composed with $H_1, \ldots, H_m$*, written $G \circ (H_1, \ldots, H_m)$ or, when delimiters are needed, $[G \circ (H_1, \ldots, H_m)]$, as the n'-ary relation *being $x_1, \ldots, x_n$ and y such that for some $z_1, \ldots, z_m$, $x_1, \ldots, x_n$ bear H_1 to z_1 and $\ldots$ and $x_1, \ldots, x_n$ bear H_m to z_m and $z_1, \ldots, z_m$ bear G to y*:

(.2) $G \circ (H_1, \ldots, H_m) =_{df}$
$[\lambda x_1 \ldots x_n y\, \exists z_1 \ldots \exists z_m (H_1 x_1 \ldots x_n z_1 \;\&\; \ldots \;\&\; H_m x_1 \ldots x_n z_m \;\&\; Gz_1 \ldots z_m y)]$

When $m = n = 1$, then (.2) reduces to (.1). Clearly, the definientia in (.1) and (.2) are core λ-expressions and so by axiom (39.2), the definientia are significant. So $G \circ H\downarrow$ and $G \circ (H_1, \dots, H_m)\downarrow$, by the theory of definitions and identity.

(**901**) **Theorems:** Operations Defined by Composition. We now establish that (.1) if H and G are unary operations, then $G \circ H$ is a unary operation such that for any x, $[G \circ H](x) = G(H(x))$; and (.2) if $H_1, \dots, H_m$ are n-ary operations ($n \geq 0$) and G is an m-ary operation ($m \geq 1$), then G composed with $H_1, \dots, H_m$, when applied to $x_1, \dots, x_n$ is identical to the result of applying G to the arguments $H_1(x_1, \dots, x_n), \dots, H_m(x_1, \dots, x_n)$:

(.1) $Op^1(H)\ \&\ Op^1(G) \rightarrow (Op^1(G \circ H)\ \&\ \forall x([G \circ H](x) = G(H(x))))$

(.2) $(Op^n(H_1)\ \&\ \dots\ \&\ Op^n(H_m)\ \&\ Op^m(G)) \rightarrow (Op^n(G \circ (H_1, \dots, H_m))\ \&$
$\forall x_1 \dots \forall x_n([G \circ (H_1, \dots, H_m)](x_1, \dots, x_n) = G(H_1(x_1, \dots, x_n), \dots, H_m(x_1, \dots, x_n))))$

(**902**) **Remark:** Rigid Restricted Variables for Numerical Operations. Note that $Op^n(R)$ is a condition with a single free variable, that it is strictly non-empty (i.e., $\vdash_\Box \exists R(Op^n(R))$, for every n, $n \geq 0$), and has strict existential import (i.e., $\vdash_\Box Op^n(\Pi) \rightarrow \Pi\downarrow$, for any n, $n \geq 0$). So $Op^n(R)$ is a restriction condition on relations, as defined in (336). Moreover, since (895) is a modally strict theorem, $Op^n(R)$ is a rigid restriction condition, as defined in (340). Consequently, in what follows, we use the undecorated variables $f, g, h, \dots$ as rigid restricted variables ranging over numerical operations. We first make use of these restricted variables in (907). But when we do, it is important to remember that $f, g, h, \dots$ can be used either as function symbols, as in $f(x_1, \dots, x_n)$, or as relation symbols, as in $f x_1 \dots x_n y$.

14.8.5 Recursively-Defined Relations and Functions

(**903**) **Remark:** How to Justify Recursive Definitions. Recursive definitions of new numerical operations are distinctive in that the symbol being defined occurs on both sides of the recursive clause of the definition. For example, the standard recursive definition of *addition on the natural numbers* is given as follows:

(.1) $n + 0 = n$
$n + m' = (n + m)'$

In this definition, the new symbol + occurs in *both* the definiendum and definiens of the recursive clause.[383]

[383]For some readers, it may be helpful to point out that recursive definitions are somewhat analogous to differential equations, which defines constraints that any function serving as the solution to the equation must satisfy.

The theory of definitions developed thus far makes no allowance for such definitions. So, if we want to introduce a definition by recursion like the above, we have to show that (a) the definition isn't circular, (b) the new symbol introduced is significant (i.e., denotes a relation), and (c) the relation signified by the new symbol is an operation on the natural numbers. All three of these goals can be accomplished by a Recursion Theorem, for such a theorem would guarantee that given any recursive definition, there is a relation which is a numerical operation that satisfies both clauses of the definition.

But it isn't immediately obvious how to prove such a Recursion Theorem, since there are only a few ways to show that relations exist, namely, appeal to either the safe extension axiom (49), the comprehension principle for relations (191), or the Kirchner Theorem (271). *None* of these ways make it obvious why the definition of a new symbol by recursion is justified.

In most logic and mathematics texts, a background of set theory is presupposed, so that the Recursion Theorem can be proved in a standard way. But in the present work, functions are defined as relations (not sets), and so a binary function symbol like +, introduced by the above recursive definition, must denote a tertiary relation. So if +, as given in (.1), is to be parsed as a new relation term, then there should be a proof that there is a relation R such that necessarily, R is exemplified by all and only numbers n, m, o such $n+m=o$.

To see the issue more clearly, let's make the relational character of (.1) explicit, by replacing the binary function symbol + with a tertiary relation symbol for *addition*, written $\boldsymbol{A}$. Let us temporarily assume that $Op^2(\boldsymbol{A})$, i.e., that *addition* is a binary numerical operation. (This will be proved later.) Then, if we use relational notation for $\boldsymbol{A}$ instead of the infix notation for +, the relational character of (.1) is made explicit by expressing it as the following two-clause stipulation:

(.2) $\boldsymbol{A}n0n$
$\boldsymbol{A}nm'(\imath k \boldsymbol{A}nmk)'$

In other words, (.2) stipulates that the binary numerical operation of *addition* relates n and Zero to n and relates n and m' to the successor of the number that results by adding n and m. Note that (.2) is not a definition of $\boldsymbol{A}$, it merely gives a pair of constraints that $\boldsymbol{A}$ must satisfy in order to correspond to +.

Now by (899.2) and (893), we can rewrite the first clause of (.2) as $\boldsymbol{A}(n,0) = n$, and if we then also make use of the operation of *successor* (896), we can rewrite the second clause of (.2) as $\boldsymbol{A}(n,m') = \boldsymbol{s}(\boldsymbol{A}(n,m))$. So to justify recursive addition, we have to show that $\boldsymbol{A}$ exists and is a numerical operation such that:

(.3) $\boldsymbol{A}(n,0) = n$
$\boldsymbol{A}(n,m') = \boldsymbol{s}(\boldsymbol{A}(n,m))$

Our plan for justifying the recursive definition of addition is to construct a new relation and prove that it is a numerical operation satisfying (.3). We will then follow the same strategy when we consider recursive definitions for n-ary functions where $n \geq 2$.

(**904**) **Remark:** Digression on Alternatives. We might try to avoid the task of proving a Recursion Theorem by regarding recursive definitions of new functions as mere metalinguistic abbreviations of complex terms of the language. For example, consider (903.1). Both the base clause and the recursive clause might be given a metalinguistic interpretation so as to avoid interpreting '+' as a term. The base clause would tell us metalinguistically that we can regard the expression $n+0$ as an abbreviation of the expression n. The recursive clause of the definition would tell us, metalinguistically, that the expression $n+m'$ abbreviates the expression $(n+m)'$. If m is not a symbol denoting Zero, then we could reapply this understanding: since every positive integer is the successor of a natural number (843), then the expression m could be regarded as the abbreviation of some expression, say k', where the expression k denotes the predecessor of the number denoted by m. So $(n+m)'$ would become an abbreviation of $(n+k')'$, which then becomes an abbreviation of $(n+k)''$. As long as the expression k doesn't denote Zero, we could continue this procedure until we reach the expression $(n+0)'^{\cdots\prime}$, for some finite sequence of primes. At this point, the base clause of the recursive definition allows us to regard the expression $n+0$ as an abbreviation of the expression n and start reducing (and eventually eliminating) the sequence of primes using the definition of the numerals (825).

But this procedure abandons the analysis of recursive functions as *relations* and, hence, the analysis of + as an operation. If we want to regard recursive definitions like (903.1) as introducing relations, we must show how such definitions yield proofs of the existence of the relations in question by one of the theorems asserting relation existence. Without forging this connection between recursive definitions and an existence theorem, recursive definitions at best introduce functional conditions rather than functional relations. Moreover, the recursion theorem we prove here is to be contrasted with those proved in Dedekind 1888 (Theorem 126), Frege 1903a (Theorem 256), Enderton 1972 [2001, 39], etc., since these all take functions as primitive and assume that functions are extensional entities (and so identical when they map the same arguments to the same values).

There is a second alternative to proving a Recursion Theorem. We could regard (903.1) as asserting two new *axioms*. When mathematicians formulate Peano Arithmetic, they often assert the recursive definition of addition as axioms. Some, for example, take Zero (0), *being a number* (N), the *successor* function ($'$), along with the operations + and $\times$, as primitive notions and then assert the Dedekind/Peano postulates along with the recursive definitions of + and × as axioms (they also assume a comprehension scheme for properties of numbers). The axioms for + and × have the same form as the recursive definitions for + and × but have a different status within Peano Arithmetic; they aren't eliminable in the way that definitions are. Indeed, with these additional axioms, Peano Arithmetic becomes a non-conservative extension of the Dedekind/Peano axioms for number theory.

The advantange of proving a Recursion Theorem in the way outlined in Remark (903) now becomes clear. The Recursion Theorem would show that recursive definitions are simply *conservative* extensions of object theory. Such definitions don't allow us to prove any new theorems not already provable in the language of the theory without those definitions.

(**905**) **Theorems and Definitions:** The Relation of Addition. Before we prove the Recursion Theorem for defining n-ary recursive functions, we show, as an example, how to inductively define a relation of addition that (a) exists by comprehension, (b) is a numerical operation, and (c) satisfies the recursive definition in (903.1). To do this, we define by induction a tertiary relation of *addition* that clearly exists by comprehension. In this and the next item, our goal is to develop the technique we'll use to prove the recursion theorem for n-ary functions. Later, once we have proved the recursion theorem for n-ary functions, we will be able to re-prove the existence of the operation of addition in object theory using the traditional recursive definition. But for now it is of interest to see that there exists by comprehension a relation that accomplishes (a), (b), and (c) – these will be established, respectively, in (.4), (.5), and (.6.a), (.6.b) below.

However, in order to define the relation of *addition*, we must first define an inductive sequence of relations and show that they are operations. We first define (.1.a) $\boldsymbol{A}_0$, i.e., *addition-by*-0. Then we prove (.1.b) that it is a unary operation. Then we show (.2.a) that if $\boldsymbol{A}_m$ is an operation then a certain composition is also an operation. Then (.2.b) we use the latter to define $\boldsymbol{A}_{m'}$, i.e, *addition-by-*m'. This yields a sequence of tertiary relations *addition-by-m* (.3) all of which are unary operations.

We therefore begin by defining $\boldsymbol{A}_0$ as the unary projection function that maps every argument to itself and prove that it is a unary numerical operation:

(.1) Base Definition of Addition-by-0 and Theorem:

(.a) $\boldsymbol{A}_0 =_{df} \pi^2_1$

(.b) $Op^1(\boldsymbol{A}_0)$

We then establish that the composition of the successor function and A_m, for any number m, is a numerical operation if A_m exists, and then define $\boldsymbol{A}_{m'}$ to be this composed function:

(.2) Inductive Theorem and Definition of Addition-by-m':

(.a) $Op^1(\boldsymbol{A}_m) \to Op^1(\boldsymbol{s} \circ \boldsymbol{A}_m)$

(.b) $\boldsymbol{A}_{m'} =_{df} \boldsymbol{s} \circ \boldsymbol{A}_m$

It now follows that for every natural number m, addition-by-m is a unary numerical operation:

(.3) $\forall m Op^1(\boldsymbol{A}_m)$

We now define the relation of *addition* as the relation *being natural numbers n, m, and j such that (the value of) addition-by-m, when applied to (the argument) n, is number-identical to j*:

(.4) $\boldsymbol{A} =_{df} [\lambda nmj\, \boldsymbol{A}_m(n) \doteq j]$

We now have the following:

(.5) $Op^2(\boldsymbol{A})$

Finally, to see that $\boldsymbol{A}$ satisfies the form of the recursive definition of addition (903.3), we show that $\boldsymbol{A}$ relates n and Zero to n; and $\boldsymbol{A}$ relates n and the successor of m to the the successor of the $\boldsymbol{A}$ of n and m:

(.6) Addition Satisfies the Recursive Definition

(.a) $\boldsymbol{A}(n,0) = n$

(.b) $\boldsymbol{A}(n,m') = \boldsymbol{s}(\boldsymbol{A}(n,m))$

Note that the successor of $\boldsymbol{A}(n,m)$, i.e., $\boldsymbol{s}(\boldsymbol{A}(n,m)) = [\boldsymbol{s} \circ \boldsymbol{A}](n,m)$.

(**906**) **Remark:** The Traditional Recursive Definition of Addition. To see that addition as defined above satisfies the form of the recursive definition of addition (903.1), note that we are justified, by (905.5), in rewriting (905.6.a) and (905.6.b) in infix and prime notation, where + replaces $\boldsymbol{A}$ and n' replaces $\boldsymbol{s}(n)$:

$$\begin{array}{lrcl} (\xi) & n+0 & = & n \\ (\zeta) & n+m' & = & (n+m)' \end{array}$$

Note that using this alternative notation, we can prove:

$$n+1 = n'$$

For as an instance of (ζ), we know:

$$n+0' = (n+0)'$$

So, using (ξ) to substitute n for $n+0$ in the above it follows that:

(ϑ) $n + 0' = n'$

But by (825), $1 = 0'$. From this and (ϑ) it follows that $n + 1 = n'$.

(**907**) **Definitions:** N-ary Numerical Recursive Definitions. To justify recursive definitions of numerical operations (i.e., to show that such definitions introduce new relations), we first have to describe their form in more general terms. In a recursive definition, a new operation f is defined in terms of some known operation and f itself.

In the case of a unary numerical operation f, the base clause stipulates that the value of f for the argument Zero is some given natural number, say i, and the recursive clause stipulates that the value of f for the argument m' is the value of a binary numerical operation g for the arguments m and $f(m)$, i.e.,

(.1) $$\begin{aligned} f(0) &= i \\ f(m') &= g(m, f(m)) \end{aligned}$$

In what follows, we plan to show how this definition is preserved in our system, where numerical functions are defined as relations. So, our plan will be to show that we can recursively define a binary relation F if given a ternary relation G such that both F and G are numerical operations. We do this by appealing to the composition of G with π^2_1 (i.e., the number-identity projection function or $\doteq$) and F, as defined in (900.2), and recasting the above in functional notation defined in terms of relation-argument structure. Thus, while the form of such a recursive definition will be:

(.2) $$\begin{aligned} &F(0) = i \\ &F(m') = [G \circ (\pi^2_1, F)](m) \end{aligned}$$

this definition will tell us that the recursively defined relation F relates Zero to i and relates m' to the result of applying the composition of G with number-identity and F to m. The definition of function application will guarantee that (.2) can be rewritten as (.1).

In the case of a *binary* numerical operation f, the traditional base clause stipulates that the value of f for the arguments n and Zero is to be the value of some unary numerical operation h for the argument n, and the recursive clause stipulates that the value of f for the arguments n and m' is the value of some ternary numerical operation g for the arguments n, m, and $f(n,m)$, i.e.,[384]

(.3) $$\begin{aligned} f(n,0) &= h(n) \\ f(n,m') &= g(n,m,f(n,m)) \end{aligned}$$

[384] Consider the binary numerical operation of addition. If we let $H(x) = x$ and $G(x,y,z) = \boldsymbol{s}(z)$, then we have $F(n,0) = H(n) = n$. And $F(n,m') = G(n,m,F(n,m)) = \boldsymbol{s}(F(n,m))$. Or, in more familiar notation $n + 0 = n$ and $n + m' = (n + m)'$.

In relational terms, we recursively define a ternary relation F if given a binary relation H and a quarternary relation G both of which are numerical operations. In the recursive clause, we appeal to the composition of G with three functions: π^3_1, π^3_2, and F. Such compositions were defined in (900.2). Thus, the above definition in functional notation is implemented in terms of relation-argument structure. The base clause of the recursive definition of F stipulates that F relates n and Zero to the natural number to which H relates n. The recursive clause states F relates n and m' to the following number: the value of the composed function $G \circ (\pi^3_1, \pi^3_2, F)$ for the arguments n and m:

(.4) $F(n, 0) = H(n)$
$F(n, m') = [G \circ (\pi^3_1, \pi^3_2, F)](n, m)$

When $G \circ (\pi^3_1, \pi^3_2, F)$ is applied to n and m, there are two steps to the evaluation. In the first step, π^3_1 takes both arguments and yields n, π^3_2 takes both arguments and yields m, and F takes both arguments and yields $F(n, m)$. In the second step, the values output in the first step are the three arguments to G so that $F(n, m')$ ends up being $G(n, m, F(n, m))$. The definition of function application will guarantee that (.3) can be rewritten as (.4).

Finally, in the case of an i'-ary numerical operation f, the base clause stipulates that the value of f for the arguments $n_1, \ldots, n_i$ and Zero is the value of an n-ary numerical operation h for the arguments $n_1, \ldots, n_i$, and the recursive clause stipulates that the value of f, for the arguments $n_1, \ldots, n_i$, and m', is the value of the i''-ary numerical operation g for the arguments $n_1, \ldots, n_i, m$, and $f(n_1, \ldots, n_i, m)$, i.e.,

(.5) $f(n_1, \ldots, n_i, 0) \;=\; h(n_1, \ldots, n_i)$
$f(n_1, \ldots, n_i, m') \;=\; g(n_1, \ldots, n_i, m, f(n_1, \ldots, n_i, m))$

To state this in relational terms, we recursively define F as an $i+2$-ary relation if given an $i1$-ary relation H and an $i+3$-ary relation G both of which are numerical operations. In the recursive clause, we appeal to the composition of G with $i+2$ functions: $\pi^{i''}_1, \ldots, \pi^{i''}_i, \pi^{i''}_{i'}$ and F. Such compositions were defined in (900.2). Again, this implements the definition given in functional notation in terms of relation-argument structure. The recursive clause states F relates $n_1, \ldots, n_i$ and m' to the following number: the value of the composed function $G \circ (\pi^{i''}_1, \ldots, \pi^{i''}_i, \pi^{i''}_{i'}, F)$ for the arguments $n_1, \ldots, n_i$ and m:

(.6) $F(n_1, \ldots, n_i, 0) = H(n_1, \ldots, n_i)$
$F(n_1, \ldots, n_i, m') = [G \circ (\pi^{i''}_1, \pi^{i''}_2, \ldots, \pi^{i''}_i, \pi^{i''}_{i'}, F)](n_1, \ldots, n_i, m)$

When $G \circ (\pi^{i''}_1, \pi^{i''}_2, \ldots, \pi^{i''}_i, \pi^{i''}_{i'}, F)$ is applied to $n_1, \ldots, n_i$ and m, there are two steps in the evaluation. In the first step, $\pi^{i''}_1$ takes the arguments $n_1, \ldots, n_i, m$ and yields the value n_1, …, $\pi^{i''}_{i'}$ takes the arguments $n_1, \ldots, n_i, m$ and yields the value m, while F takes the arguments $n_1, \ldots, n_i, m$ and yields the value $F(n_1, \ldots, n_i, m)$. In the second step, the values output by the first step are the arguments to G so that $F(n_1, \ldots, n_i, m')$ becomes $G(n_1, \ldots, n_i, m, F(n_1, \ldots, n_i, m))$. So the definition of function application guarantees that (.5) can be rewritten as (.6).

In what follows, we focus first on (.3) and (.4) and then turn to (.5) and (.6).

(**908**) **Theorems and Definitions:** Lemmas for the Recursion Theorem. Our strategy is to construct a binary numerical operation $\boldsymbol{F}$ as follows:

- start with a given unary numerical operation H and a ternary numerical operation G,
- define an inductive sequence of binary relations $\boldsymbol{F}_m$ relative to H and G,
- show, by induction, that each $\boldsymbol{F}_m$ is a unary numerical operation,
- define the tertiary relation $\boldsymbol{F}$, relative to H and G, in terms of the sequence of relations $\boldsymbol{F}_m$,
- show that $\boldsymbol{F}$ is a binary numerical operation, and
- show that $\boldsymbol{F}$ satisfies the conditions of a numerical operation recursively defined in terms of H and G.

In the definitions that follows, we *suppress*, for readability, the indices that relativize the relations $\boldsymbol{F}_m$ and the relation $\boldsymbol{F}$ to H and G. But we introduce the indices later, since they are needed to state the Recursion Theorem.

We first define (.1.a) $\boldsymbol{F}_0$ to be H. So clearly (.1.b) $\boldsymbol{F}_0$ is a unary numerical operation given that H is. We then show (.2.a) that given any m, if $\boldsymbol{F}_m$ is a unary numerical operation, then the composition of G with three unary numerical operations, namely the projection function π^2_1, the constant function $\mathcal{C}^2_m$, and $\boldsymbol{F}_m$, is also a unary numerical operation. Then we define $\boldsymbol{F}_{m'}$ as that composition.

(.1) Base Definition of $\boldsymbol{F}_0$ and Theorem:

(.a) $\boldsymbol{F}_0 =_{df} H$

(.b) $Op^1(\boldsymbol{F}_0)$

(.2) Theorem For, and Definition of, the Recursive Numerical Operation $\boldsymbol{F}_{m'}$:

(.a) $Op^1(\boldsymbol{F}_m) \rightarrow Op^1(G \circ (\pi^2_1, \mathcal{C}^2_m, \boldsymbol{F}_m))$

(.b) $\boldsymbol{F}_{m'} =_{df} G \circ (\pi^2_1, \mathcal{C}^2_m, \boldsymbol{F}_m)$

At this point it should be clear that $\boldsymbol{F}_m$ has been defined relative to H and G, and that strictly speaking, should be indexed to these initial relations.

It now follows that (.3) for every natural number m, the $\boldsymbol{F}$-operation-by-m is a unary numerical operation:

(.3) $\forall m Op^1(\boldsymbol{F}_m)$

We next define (.4) $\boldsymbol{F}$ as the relation *being natural numbers n, m, and j such that* (*the value of*) *the $\boldsymbol{F}$-operation-by-m, when applied to* (*the argument*) *n, is j*:

(.4) $\boldsymbol{F} =_{df} [\lambda nmj\, \boldsymbol{F}_m(n) \doteq j]$

Strictly speaking, $\boldsymbol{F}$ has been defined relative to H and G, but for now, we are suppressing the indices.

It now follows that (.5) $\boldsymbol{F}$ is a binary numerical operation:

(.5) $Op^2(\boldsymbol{F})$

Finally, note that $\boldsymbol{F}$ satisfies the conditions of a binary numerical operation recursively defined in terms of H and G:

(.6) $\boldsymbol{F}$ Satisfies the Recursive Definition Conditions:

(.a) $\boldsymbol{F}(n, 0) = H(n)$

(.b) $\boldsymbol{F}(n, m') = G(n, m, \boldsymbol{F}(n, m))$

Note that $G(n, m, \boldsymbol{F}(n, m)) = [G \circ (\pi^3_1, \pi^3_2, \boldsymbol{F})](n, m)$.

(**909**) **Theorem:** The Recursion Theorem for Recursive Binary Numerical Operations. Let H now be a (rigid) restricted variable ranging over unary numerical operations, and let G be a (rigid) restricted variable ranging over ternary numerical operations. These are rigid restricted variables because anything that is a numerical operation is necessarily so. Since the relation $\boldsymbol{F}$ defined in (908.4) was, strictly speaking, defined in terms of a given H and G, we shall henceforth write $\boldsymbol{F}$ as $\boldsymbol{F}_{H,G}$. Then the recursion theorem asserts that $\boldsymbol{F}_{H,G}$ is a binary numerical operation that satisfies the conditions of recursion:

$$Op^2(\boldsymbol{F}_{H,G}) \,\&\, \boldsymbol{F}_{H,G}(n, 0) = H(n) \,\&\, \boldsymbol{F}_{H,G}(n, m') = G(n, m, \boldsymbol{F}_{H,G}(n, m))$$

(**910**) **Theorems and Definitions:** Lemmas for the N-ary Recursion Theorem. We now construct an i'-ary numerical operation $\boldsymbol{F}$ using the following sequence of steps:

- start with an i-ary numerical operation H and an i''-ary numerical operation G ($i \geq 0$),
- define an inductive sequence of i+1-ary relations $\boldsymbol{F}_m$ relative to H and G,

- show, by induction, that each $\boldsymbol{F}_m$ is an i-ary operation,
- define the i+2-ary relation $\boldsymbol{F}$, relative to H and G, in terms of the sequence of relations,
- show that $\boldsymbol{F}$ is an i'-ary operation, and
- show that $\boldsymbol{F}$ satisfies the conditions of an operation recursively defined in terms of H and G.

Again, in the definitions that follows, we *suppress*, for readability, the indices that relativize the relations $\boldsymbol{F}_m$ and the relation $\boldsymbol{F}$ to H and G. But we introduce the indices later, since they are needed to state the Recursion Theorem.

We first define (.1.a) $\boldsymbol{F}_0$ to be H. So clearly (.1.b) $\boldsymbol{F}_0$ is an i-ary numerical operation given that H is. We then show (.2.a) that given any m, if $\boldsymbol{F}_m$ is an i-ary numerical operation, then the composition of G with i'' numerical operations each of which is i-ary, namely the projection functions $\pi_1^{i'},\ldots,\pi_i^{i'}$, the constant function $\mathcal{C}_m^{i'}$, and $\boldsymbol{F}_m$, is also a i-ary operation. Then we define $\boldsymbol{F}_{m'}$ as that composition.

(.1) Base Definition of $\boldsymbol{F}_0$ and Theorem:

(.a) $\boldsymbol{F}_0 \ =_{df}\ H$

(.b) $Op^i(\boldsymbol{F}_0)$

(.2) Theorem For, and Definition of, the Recursive Operation $\boldsymbol{F}_{m'}$:

(.a) $Op^i(\boldsymbol{F}_m) \to Op^i(G \circ (\pi_1^{i'},\ldots,\pi_i^{i'},\mathcal{C}_m^{i'},\boldsymbol{F}_m))$

(.b) $\boldsymbol{F}_{m'} =_{df} G \circ (\pi_1^{i'},\ldots,\pi_i^{i'},\mathcal{C}_m^{i'},\boldsymbol{F}_m)$

At this point it should be clear that $\boldsymbol{F}_m$ has been defined relative to H and G, and that strictly speaking, should be indexed to these initial relations. Note also that when $i = 0$, there are no projection functions and so, for the composition, G just takes the nullary functions $\mathcal{C}_m^1$ and $\boldsymbol{F}_m$ as its two arguments.

It now follows that (.3) for every natural number m, the $\boldsymbol{F}$-operation-by-m is an i-ary numerical operation:

(.3) $\forall m Op^i(\boldsymbol{F}_m)$

We now define (.4) $\boldsymbol{F}$ as the relation *being natural numbers* $n_1,\ldots,n_i,m$, *and* j *such that* (*the value of*) *the* $\boldsymbol{F}$*-operation-by-*m, *when applied to* (*the argument*) $n_1,\ldots,n_i$, *is* j:

(.4) $\boldsymbol{F} =_{df} [\lambda n_1 \ldots n_i m j\ \boldsymbol{F}_m(n_1,\ldots,n_i) \doteq j]$

Strictly speaking, $\boldsymbol{F}$ has been defined relative to H and G, but for now, we are suppressing the indices. Note also that when $i = 0$, the $\boldsymbol{F}$*-operation-by-m* is nullary and there are no arguments $n_1, \ldots, n_i$.

It now follows that $\boldsymbol{F}$ is a i'-ary numerical operation:

(.5) $Op^{i'}(\boldsymbol{F})$

Finally, $\boldsymbol{F}$ satisfies the conditions of a binary numerical operation recursively defined in terms of H and G:

(.6) $\boldsymbol{F}$ Satisfies the Recursive Definition Conditions:

(.a) $\boldsymbol{F}(n_1, \ldots, n_i, 0) = H(n_1, \ldots, n_i)$

(.b) $\boldsymbol{F}(n_1, \ldots, n_i, m') = G(n_1, \ldots, n_i, m, \boldsymbol{F}(n_1, \ldots, n_i, m))$

Note that $G(n_1, \ldots, n_i, m, \boldsymbol{F}(n_1, \ldots, n_i, m)) = [G \circ (\pi_1^{i''}, \ldots, \pi_{i+1}^{i''}, \boldsymbol{F})](n_1, \ldots, n_i, m)$.

(**911**) **Theorem:** The Recursion Theorem for N-ary Recursive Numerical Operations. Let H now be a restricted variable ranging over i-ary numerical operations, and let G be a restricted variable ranging over i''-ary numerical operations. Again, these are rigid restricted variables since anything that is a numerical operation is necessarily so. Since the relation $\boldsymbol{F}$ defined in (910.4) was, strictly speaking, defined in terms of a given H and G, we shall henceforth write $\boldsymbol{F}$ as $\boldsymbol{F}_{H,G}$. Then the recursion theorem asserts that $\boldsymbol{F}_{H,G}$ is an i'-ary numerical operation that satisfies the conditions of recursion:

$$\begin{array}{l} Op^{i'}(\boldsymbol{F}_{H,G}) \,\&\, \boldsymbol{F}_{H,G}(n_1, \ldots, n_i, 0) = H(n_1, \ldots, n_i) \,\& \\ \quad \boldsymbol{F}_{H,G}(n_1, \ldots, n_i, m') = G(n_1, \ldots, n_i, m, \boldsymbol{F}_{H,G}(n_1, \ldots, n_i, m)) \end{array}$$

(**912**) **Definitions and Theorems:** Traditional Recursive Definitions and Theorems. With our justification for recursive definitions of numerical operations firmly in hand, we may now officially employ the traditional recursive definitions to introduce other numerical operations such as addition (which we reintroduce!), multiplication, exponentiation, and factorialization on the natural numbers.

• We define Addition ($\boldsymbol{A}$) as the following binary numerical operation:

(.1.a) $\boldsymbol{A} =_{df} \boldsymbol{F}_{\pi_1^2, \boldsymbol{s} \circ \pi_3^4}$

That is, addition is the tertiary relation obtained by the Recursion Theorem (911) when H is the projection function π_1^2 and G is the composition of the successor function with the projection function π_3^4. Note that π_1^2 just is the relation $\doteq$, so the base case of the definition is the same is in (905.1.a). The recursive clause now makes use of our apparatus for defining n-ary recursive functions (911).

To recast H and G in more familiar notation, where we use x, y, and z as rigid, restricted variables ranging over the natural numbers (to avoid clash of variables below), we have $H(x) = \pi^2_1(x) = x$ and $G(x,y,z) = [\boldsymbol{s} \circ \pi^4_3](x,y,z) = \boldsymbol{s}(z) = z'$. It now follows that:

(.1.b) $\boldsymbol{A}(n,0) = n$
$\boldsymbol{A}(n,m') = (\boldsymbol{A}(n,m))'$

The proof is in the Appendix. Note how this captures the traditional recursive definition of addition; using + in infix notation instead of $\boldsymbol{A}$ in prefix notation, these become:

$$n + 0 = n$$
$$n + m' = (n + m)'$$

Henceforth, we write $\boldsymbol{A}$ when composing functions with the addition function but write + in infix notation when we're adding numbers.

• We define Multiplication ($\boldsymbol{M}$) as the following binary numerical operation:

(.2.a) $\boldsymbol{M} =_{df} \boldsymbol{F}_{\mathcal{C}^2_0, \boldsymbol{A}\circ(\pi^4_1,\pi^4_3)}$

That is, multiplication is the tertiary relation obtained by the Recursion Theorem (911) when H is the constant function $\mathcal{C}^2_0$ and G is the composition of the addition function with the two projection functions π^4_1 and π^4_3. Intuitively, $H(x) = \mathcal{C}^2_0(x) = 0$ and $G(x,y,z) = [\boldsymbol{A} \circ (\pi^4_1, \pi^4_3)](x,y,z) = \boldsymbol{A}(x,z) = x + z$. It now follows that:

(.2.b) $\boldsymbol{M}(n,0) = 0$
$\boldsymbol{M}(n,m') = n + \boldsymbol{M}(n,m)$

The proof is in the Appendix. Note how this captures the traditional recursive definition of multiplication; using $\times$ in infix notation instead of $\boldsymbol{M}$ in prefix notation, these becomes:

$$n \times 0 = 0$$
$$n \times m' = n + (n \times m)$$

Henceforth, we write $\boldsymbol{M}$ when composing functions with the multiplication function but write $\times$ in infix notation when we're multiplying numbers.

• We define Exponentiation ($\boldsymbol{E}$) as the following binary numerical operation:

(.3.a) $\boldsymbol{E} =_{df} \boldsymbol{F}_{\mathcal{C}^2_1, \boldsymbol{M}\circ(\pi^4_1,\pi^4_3)}$

That is, exponentiaion is the tertiary relation obtained by the Recursion Theorem (911) when H is the constant function $\mathcal{C}^2_1$ and G is the composition of the multiplication function with the two projection functions π^4_1 and π^4_3. Intuitively, $H(x) = \mathcal{C}^2_1 = 1$ and $G(x,y,z) = [\boldsymbol{M} \circ (\pi^4_1, \pi^4_3)](x,y,z) = \boldsymbol{M}(x,z) = x \times z$. It now follows that:

(.3.b) $\boldsymbol{E}(n,0) = 1$
$\boldsymbol{E}(n,m') = n \times \boldsymbol{E}(n,m)$

Note how this captures the traditional recursive definition of exponentiation; using the standard notation $()^{()}$ instead of $\boldsymbol{E}$ in prefix notation, these becomes:

$$\begin{aligned} n^0 &= 1 \\ n^{m'} &= n \times n^m \end{aligned}$$

Henceforth, we write $\boldsymbol{E}$ when composing functions with the exponentiation function but use the standard notation $()^{()}$ when we're raising a number to an exponent.

• We define Factorialization ($\boldsymbol{\Pi}$) as the following unary numerical operation:

(.4.a) $\boldsymbol{\Pi} =_{df} \boldsymbol{F}_{\mathcal{C}^1_1, \boldsymbol{M} \circ (\pi^3_2, \boldsymbol{s} \circ \pi^3_1)}$

That is, factorialization is the binary relation obtained by the Recursion Theorem (911) when H is the constant function $\mathcal{C}^1_1$ and G is the composition of the multiplication function with the projection function π^3_2 and the composition of the successor function with the projection function π^3_1., i.e., $\boldsymbol{M} \circ (\pi^3_2, \boldsymbol{s} \circ \pi^3_1)$. Intuitively, $H = \mathcal{C}^1_1 = 1$ and $G(y,z) = [\boldsymbol{M} \circ (\pi^3_2, \boldsymbol{s} \circ \pi^3_1)](y,z) = \boldsymbol{M}(z,y') = z \times y'$. It now follows that:

(.4.b) $\boldsymbol{\Pi}(0) = 1$
$\boldsymbol{\Pi}(n') = \boldsymbol{\Pi}(n) \times n'$

Note how this captures the traditional recursive definition of factorialization; using ! in postfix notation instead of $\boldsymbol{\Pi}$ in prefix notation, these become:

$$\begin{aligned} 0! &= 1 \\ n'! &= n! \times n' \end{aligned}$$

Henceforth, we write $\boldsymbol{\Pi}$ when composing functions with the factorialization function but write ! in postfix notation when referring to a number's factorialization.

We leave the definition of many other primitive recursive functions to the reader; they can be found in classic texts, or in reference works such as Dean 2020.

(**913**) **Remark:** Primitive Recursive vs. Recursive Functions. Since we've now (a) defined the basic or *initial* functions, namely, the successor function ($\boldsymbol{s}$), the constant functions ($\mathcal{C}^{n'}_m$), and the projection functions ($\pi^{i'}_k$), and (b) shown that we can derive new functions by composition ($G \circ (H_1, \ldots, H_m)$) and recursion ($\boldsymbol{F}_{H,G}$), we have reconstructed the entire class of *primitive recursive functions*. In order to reconstruct the class of *general* recursive functions, we need only show that the minimization operator can be defined in our system.

(**914**) **Definition:** The Minimization Function. Where f is any j+1 numerical operation ($j \geq 0$), we define the (restricted) variable-binding μ operator as follows. *The least* natural number n *such that* f maps $m_1,\ldots,m_j,n$ to Zero is the natural number n such that f maps $m_1,\ldots,m_j,n$ to Zero and for any number i less than n, $f(m_1,\ldots,m_j,i)$ is defined and equal to a value other than Zero:

$$\begin{aligned}&\mu n(f(m_1,\ldots,m_j,n)=0) =_{df}\\ &\quad \imath n(f(m_1,\ldots,m_j,n)=0 \;\&\; \forall i(i<n \rightarrow \exists k(f(m_1,\ldots,m_j,i)=k \;\&\; k\neq 0))\end{aligned}$$

Note that if there is no minimal n such that $f(m_1,\ldots,m_j,n) = 0$, or if there is some i such that $i < n$ and $\neg f(m_1,\ldots,m_j,i)\downarrow$, then the description will not be satisfiable in which case $\mu nf(m_1,\ldots,m_j,n)$ will not be significant, i.e., our logic will guarantee that $\neg\mu nf(m_1,\ldots,m_j,n)\downarrow$.

14.9 Deriving 2nd-Order Peano Arithmetic

(**915**) **Remark:** Interpreting the Language of Second-order Peano Arithmetic (PA2). We now show that the system of PA2 is derivable in object theory. To do this, we show that there is a translation of the language of PA2 into the present language for which the axioms of PA2 become derivable as theorems. We use as our basis for the language of PA2 the system described in Simpson 1999 [2009]. In Simpson 1999 [2009] (2–3), we find the the language of PA2 is described in stages. His presentation may be re-parsed as follows:

- Simple terms:
  - Number constants: 0 and 1
  - Number variables: $i,j,k,m,n,\ldots$
  - Set variables: $X,Y,Z,\ldots$ (intended to range over all subsets of ω).
- Complex number terms: whenever τ_1 and τ_2 are any number terms, the following are also number terms:
  - $\tau_1 + \tau_2$
  - $\tau_1 \cdot \tau_2$

  Here + and $\cdot$ are binary numerical operation symbols intended to denote addition and multiplication of natural numbers and the numerical terms are intended to denote natural numbers.
- Atomic formulas: Where τ_1 and τ_2 are any number terms and X is any set variable, the following are formulas:
  - $\tau_1 = \tau_2$

- $\tau_1 < \tau_2$
- $\tau_1 \in X$

The intended meanings of these respective atomic formulas are that τ_1 equals τ_2, τ_1 is less than τ_2, and τ is an element of X.

- Complex Formulas: whenever φ and ψ are formulas, n is a number variable and X is a set variable, then $\varphi \wedge \psi$, $\varphi \vee \psi$, $\neg\varphi$, $\varphi \to \psi$, $\varphi \leftrightarrow \psi$, $\forall n\varphi$, $\exists n\varphi$, $\forall X\varphi$, $\exists X\varphi$ are formulas.

Now we may translate the terms and formulas of PA2 into object theory as follows.

- Simple Terms
  - Number constants: The symbols 0 and 1 of PA2 are to be translated as the symbols 0 and 1, as defined in (782.1) and (825.1).
  - Number variables are to be translated as the rigid, restricted variables ranging over numbers, as the latter are defined in (807).
  - Set variables are to be translated as the property variables $F, G, H, \ldots$.

If we use the decorated metavariables τ_1^* and τ_2^* to designate the individual terms of object theory that serve as the translation of the PA2 number terms τ_1 and τ_2, respectively, and use φ^* and ψ^*, respectively, to designate the translations of PA2 formulas φ and ψ, then we complete the translation as follows:

- Complex Number Terms:
  - The term $\tau_1 + \tau_2$ is to be translated as $\tau_1^* + \tau_2^*$, as defined in (912.1)
  - the term $\tau_1 \cdot \tau_2$ is to be translated as $\tau_1^* \times \tau_2^*$, as defined in (912.2)
- Atomic formulas:
  - $\tau_1 = \tau_2$ is to be translated as $\tau_1^* \doteq \tau_2^*$, as defined in (839.1)
  - $\tau_1 < \tau_2$ is to be translated as $\tau_1^* < \tau_2^*$, as defined in (830.1)
  - $\tau_1 \in X$ is to be translated as $F\tau_1^*$, where F is the translation of X and $F\tau_1^*$ is an exemplification formula as defined in (3.3.a).
- Complex Formulas:
  - The formulas $\neg\varphi$ and $\varphi \to \psi$ are to be translated using the primitive connectives $\neg$ and $\to$ (2.1) and (2.2), respectively, and so are to be translated as the formulas $\neg\varphi^*$ and $\varphi^* \to \psi^*$.

– The formulas $\varphi \wedge \psi$, $\varphi \vee \psi$, and $\varphi \leftrightarrow \psi$ are to be translated into the formulas $\varphi^* \,\&\, \psi^*$, $\varphi^* \vee \psi^*$, and $\varphi^* \equiv \psi^*$, as defined in (18.1), (18.2) and (18.3), respectively.

– The formulas $\forall n\varphi$ and $\exists n\varphi$ are to be translated using both (a) the primitive quantifier $\forall$ (2.3) and the defined quantifier $\exists$ (18.4), respectively, and (b) the rigid, restricted variables ranging over numbers, so that $\forall n\varphi$ and $\exists n\varphi$ become translated as $\forall n\varphi^*$ and $\exists n\varphi^*$, respectively. (Note that we can expand these, by eliminating the restricted variables, to $\forall x(\mathbb{N}x \to \varphi^{*x}_{\;n})$ and $\exists x(\mathbb{N}x \,\&\, \varphi^{*x}_{\;n})$, respectively.)

– The formulas $\forall X\varphi$ and $\exists X\varphi$ are to be translated similarly but with the variable F replacing X (so that the quantifiers bind the property variable F), with the result that $\forall X\varphi$ and $\exists X\varphi$ become translated as $\forall F\varphi^*$ and $\exists F\varphi^*$.

Finally, Simpson states the axioms of PA2 as follows (modified only to change the order in which the variables occur in the statement, so that the statements match our convention of using n as the primary variable, then m, etc.):

(i) basic axioms:

$$
\begin{aligned}
&n+1 \neq 0\\
&n+1 = m+1 \to n=m\\
&n+0 = n\\
&n+(m+1) = (n+m)+1\\
&n \cdot 0 = 0\\
&n \cdot (m+1) = n + (n \cdot m)\\
&\neg(n<0)\\
&n < m+1 \leftrightarrow (n<m \vee n=m)
\end{aligned}
$$

(ii) Induction Axiom:

$$(0 \in X \wedge \forall n(n \in X \to n+1 \in X)) \to \forall n(n \in X)$$

(iii) Comprehension Scheme:

$\exists X\forall n(n \in X \leftrightarrow \varphi)$, where φ is any formula of the language of PA2 in which X doesn't occur free.

When the translation scheme introduced above is expanded to include the translation of of $n+1$ as n', as this latter was defined in (821), the above axioms of PA2 become translated as:

(i) basic axioms:

$$n' \neq 0$$
$$n' = m' \to n = m$$
$$n + 0 = n$$
$$n + m' = (n + m)'$$
$$n \times 0 = 0$$
$$n \times m' = n + (n \times m)$$
$$\neg(n < 0)$$
$$n < m' \equiv (n < m \vee n = m)$$

(ii) Induction Axiom:

$$(F0 \,\&\, \forall n(Fn \to Fn')) \to \forall n Fn)$$

(iii) Comprehension Scheme:

$\exists F \forall n(Fn \equiv \varphi^*)$, where φ^* is the translation of φ and F doesn't occur free in φ^*.

In light of the foregoing, we now have the following theorem.

(**916**) **Theorem:** The Axioms of PA2 Are Theorems. A proof is in the Appendix. So a significant fragment of mathematics is derivable in object theory. Cf. Hilbert & Bernays 1934–1939 and Simpson 1999 [2009].

14.10 Infinity

(**917**) **Definition:** Finite and Infinite Cardinals. We defined natural cardinals in (777) and established that $NaturalCardinal(x) \to \Box NaturalCardinal(x)$ (779.1). So we may introduce κ as a rigid restricted variable ranging over natural cardinals. We now say (.1) κ is *finite* if and only if κ is a natural number, and (.2) κ is *infinite* if and only if κ is not finite:

(.1) $Finite(\kappa) \equiv_{df} \mathbb{N}\kappa$

(.2) $Infinite(\kappa) \equiv_{df} \neg Finite(\kappa)$

These definitions follow Frege's conception of an infinite number and we examine the connections later in Remark (921).

(**918**) **Theorem:** There Exists an Infinite Cardinal. It is a straightforward consequence of previous definitions and theorems that (.1) any object that bears the weak ancestral of *predecessor* to a natural number is itself a natural number; (.2) no natural number numbers the natural numbers; (.3) the number of the property *being a natural number* is infinite; (.4) the number of $\mathbb{N}$ is a natural cardinal; and (.5) there is an infinite cardinal:

(.1) $\forall x(\mathbb{P}^{+}xm \to \mathbb{N}x)$

(.2) $\neg\exists n Numbers(n, \mathbb{N})$

(.3) $Infinite(\#\mathbb{N})$

(.4) $NaturalCardinal(\#\mathbb{N})$

(.5) $\exists\kappa Infinite(\kappa)$

We have formulated (.1) as a universal claim instead of asserting the open formula $\mathbb{P}^{+}xm \to \mathbb{N}x$ because the proof is by induction on m: we want the inductive hypothesis to be a universal claim and not a claim about an arbitrary, but fixed, object x.

The proof of (.5) establishes that the existence of an infinite cardinal has been derived in a system with no mathematical primitives! We've not used any primitive notions, or asserted any axioms, from number theory or set theory to define and prove there is an infinite number.

(**919**) **Definition:** Aleph$_0$. Since we've shown that $\#\mathbb{N}$ is infinite and a natural cardinal, we may introduce a name for it appropriate to its being an infinite natural cardinal. We define $\aleph_0$ to be the number of *being a natural number*:

$$\aleph_0 =_{df} \#\mathbb{N}$$

(**920**) **Theorems:** Facts about $\aleph_0$. The preceding definition yields the following immediate consequences. (.1) $\aleph_0$ is a natural cardinal; (.2) $\aleph_0$ is infinite; (.3) $\aleph_0$ is not finite; (.4) $\aleph_0$ is not a natural number; (.5) $\aleph_0$ isn't a member of the predecessor series beginning wth Zero; (.6) no natural number immediately precedes $\aleph_0$; and (.7) $\aleph_0$ immediately precedes itself:

(.1) $NaturalCardinal(\aleph_0)$

(.2) $Infinite(\aleph_0)$

(.3) $\neg Finite(\aleph_0)$

(.4) $\neg\mathbb{N}\aleph_0$

(.5) $\neg\mathbb{P}^{+}0\aleph_0$

(.6) $\neg\exists n\mathbb{P}n\aleph_0$

(.7) $\mathbb{P}\aleph_0\aleph_0$

Note that (.7) demonstrates why we didn't define $<$ as $\mathbb{P}^{*}$ (and $\leq$ as $\mathbb{P}^{+}$), for since $\mathbb{P}xy$ implies $\mathbb{P}^{*}xy$ (789.1), (.7) would have then implied the unintuitive claim $\aleph_0 < \aleph_0$ under such a definition.

It is also worth observing that (.5) corresponds to Frege's Theorem 167 (1893, §125 [2013, 154]), which asserts $\vdash \not{0} \frown (\overset{\frown}{\omega} \frown \acute{\smile}\mathfrak{f})$. We discuss this point further in the following Remark.

(**921**) **Remark:** On Frege's Definition of Finite and Infinite Cardinals. By theorem (807.3) we know $\mathbb{N}x \equiv \mathbb{P}^{+}0x$. But in a number of places, Frege indicated that $\mathbb{P}^{+}0x$ is the definiens for the notion x is a *finite cardinal number*.

In Frege 1884, §83 is titled 'Definition of Finite Number' and in this section, Frege writes:

> . . . I define as follows: the proposition "n is a member of the series of natural numbers beginning with 0" is to mean the same as "n is a finite number."

In 1893, §46 [2013, 60], Frege writes:[385]

> Accordingly, $\Delta \frown (\Theta \frown \acute{\smile}\Upsilon)$ is the truth-value of: that Θ belongs to the Υ-series starting with Δ. Thus, $\not{0} \frown (\Theta \frown \acute{\smile}\mathfrak{f})$ is the truth-value of: that Θ belongs to the cardinal number series starting with $\not{0}$, for which I can also say Θ is a *finite* cardinal number.

If we substitute 0 for $\not{0}$, x for Θ, and $\mathbb{P}^{+}$ for $\acute{\smile}\mathfrak{f}$, then the statement $\not{0} \frown (\Theta \frown \acute{\smile}\mathfrak{f})$ would be written in our notation as $\mathbb{P}^{+}0x$, which in effect is the definiens for $\mathbb{N}x$ in (807).

Furthermore, in 1893, the formula to be proved in Section Zeta (Z) [2013, 137] is:

$$\vdash\!\!\begin{array}{l} \top\, b \frown (b \frown \underline{\acute{\ }}\mathfrak{f}) \\ \llcorner\, \not{0} \frown (b \frown \acute{\smile}\mathfrak{f}) \end{array}$$

Frege then writes:

> The proposition mentioned in the main heading states that no object belonging to the cardinal number series starting with Zero follows after itself in the cardinal number series. Instead, we could also say: "No *finite* cardinal number follows after itself in the cardinal number series".

The formula in question becomes proved as Theorem 145 (2013, 144). In our notation, substituting appropriately, Frege's formula would be written as:

$$\mathbb{P}^{+}0x \rightarrow \neg\mathbb{P}^{*}xx$$

[385]In the following quote, we've corrected a known one-character transcription error in the first edition of the Ebert and Rossberg 2013 translation, p. 60. In the second sentence of the following quotation, the character $\mathfrak{f}$ has been substituted for Υ in the formula. This correction is based on the original (1893, p. 60) and is included in the revised, paperback edition of the translation (2016).

and in (838.1), we proved this in the form $n \not< n$.

Moreover, in Frege 1893, the formula to be proved in Section Iota (I) [2013, 150] is:

$$\vdash \emptyset \frown (\infty \frown \smile\!\mathfrak{f})$$

To explain this claim, Frege writes (1893, §122 [2013, 150]):

> There are cardinal numbers that do not belong to the cardinal numbers series beginning with $\emptyset$, or, as we shall also say, that are not finite, that are infinite. One such cardinal number is that of the concept *finite cardinal number*; I propose to call it *Endlos* and designate it with '∞'. I define it thus:
>
> $$\Vdash \wp(\emptyset \frown \not{\mathfrak{F}} \smile\!\mathfrak{f}) = \infty \qquad \text{(M}$$
>
> For $\emptyset \frown \not{\mathfrak{F}} \smile\!\mathfrak{f}$ is the extension of the concept *finite cardinal number*. The proposition mentioned in the heading says that the cardinal number Endlos is not a finite cardinal number.

So the formula to be proved in Section Iota asserts that Zero doesn't bear the weak ancestral of the predecessor relation to *Endlos*. This is established as Theorem 167 (2013, 154). Similarly, in object theory, there is a distinguished abstract object that numbers the natural (finite) numbers, namely $\#\mathbb{N}$, and it is provably an infinite cardinal, by (918.4) and (918.5). The traditional name that we've introduced in (919) for the number of natural numbers, $\aleph_0$, clearly corresponds to Frege's name ∞. Thus, the claim in the main heading of Frege's Section Iota becomes, in our notation, $\neg\mathbb{P}^{+}0\aleph_0$.

Finally, in 1893, the formula to be proved in Section Kappa (K) [2013, 201] is:

$$\begin{array}{l} \vdash\!\!\!\!\underline{\qquad\qquad\quad} \emptyset \frown (\wp u \frown \smile\!\mathfrak{f}) \\ \quad \llcorner\!\!\top\!\!\smile^{\mathfrak{A}}\!\!\smile^{\mathfrak{q}}\!\!\top\; u = \mathfrak{A} \,\underline{\measuredangle}\, \mathfrak{q} \end{array}$$

Frege then explicates this by writing:

> For finite cardinal numbers we can prove a proposition similar to the last, namely that the cardinal number of a concept is finite if the objects falling under it can be ordered into a *simple* (non-branching, non-looping back into itself) series starting with a certain object and ending with a certain object.

The formula to be proved ultimately becomes Theorem 327 (2013, 224). It is a conditional and in the consequent of this conditional, Frege continues to use $\emptyset \frown (\wp u \frown \smile\!\mathfrak{f})$, i.e., $\wp u$ is a member of the predecessor series starting with $\emptyset$, to express that the number of u ($\wp u$) is finite.

Given these examples, it is clear that Frege regards a cardinal number κ as *finite* just in case Zero bears the weak ancestral of *predecessor* to κ, i.e., just in case κ is a natural number. And just as clearly, he suggests that a cardinal is *infinite* just in case it fails to be finite. So our definitions in (917) conform with Frege's usage.

(**922**) **Definition:** Infinite Classes. Using the definition of $ClassOf(x,G)$ (312.1), we now stipulate that (.1) x is an *infinite class of* Gs if and only if x is a class of Gs that is numbered by an infinite cardinal, and that (.2) x is an *infinite class* if and only if for some property G, x is an infinite class of Gs:

(.1) $InfiniteClassOf(x,G) \equiv_{df} G{\downarrow} \;\&\; ClassOf(x,G) \;\&\; \exists\kappa(Infinite(\kappa) \;\&\; Numbers(\kappa,G))$

(.2) $InfiniteClass(x) \equiv_{df} \exists G(InfiniteClassOf(x,G))$

(**923**) **Theorem:** Existence of An Infinite Class. We can now prove that there exists an infinite class:

$$\exists x InfiniteClass(x)$$

So, we've established the existence of an infinite class without appealing to any mathematical primitives or asserting any mathematical axioms. The proof can be given by modally strict means because, intuitively, at every world w, the extension-at-w of $\mathbb{N}$ (i.e., $\epsilon_w\mathbb{N}$), as defined in (565), is infinite.

Of course, we could have identified $\epsilon\mathbb{N}$ as an infinite class, and although it provably is one, the proof is not modally strict.[386] The claim $InfiniteClass(\epsilon\mathbb{N})$ is a theorem but not modally strict because $\epsilon\mathbb{N}$ is not necessarily a class; this is a consequence of (320.5), which tells us no class is necessarily a class! Of course, if you use the notion of an *actual extension* (i.e., an *actual class*), as described in (335), one could show that the *actual* extension of $\mathbb{N}$ is an infinite class, by modally strict means.

[386] By definition (922.2), we have to show $\exists G(InfiniteClassOf(\epsilon\mathbb{N},G))$. But clearly, our witness is $\mathbb{N}$, so it suffices to show $InfiniteClassOf(\epsilon\mathbb{N},\mathbb{N})$, i.e., by (922.1) that:

$$\mathbb{N}{\downarrow} \;\&\; ClassOf(\epsilon\mathbb{N},\mathbb{N}) \;\&\; \exists\kappa(Infinite(\kappa) \;\&\; Numbers(\kappa,\mathbb{N}))$$

Clearly, $\mathbb{N}{\downarrow}$. By (327)★ and definition (312.1), $ClassOf(\epsilon\mathbb{N},\mathbb{N})$. Since we know $\#\mathbb{N}$ is a natural cardinal (918.5), it remains to show $\#\mathbb{N}$ is our witness to $\exists\kappa(Infinite(\kappa) \;\&\; Numbers(\kappa,\mathbb{N}))$. But $Infinite(\#\mathbb{N})$, by (918.4). And we have $Numbers(\#\mathbb{N},\mathbb{N})$ by (774.5) and the rigidity of $\mathbb{N}$ (809.2).

Chapter 15

Typed Object Theory and its Applications

In this chapter, we develop the type-theoretic version of object theory and apply it in various ways. Many of these applications consist of object-theoretic analyses of contexts and constructions in natural language. The most notable application of typed object theory is the analysis of theoretical mathematical language, for it constitutes an object-theoretic *reduction* of theoretical mathematical objects and theoretical mathematical relations.[387] We shall not only identify mathematical objects such as the real numbers and the Zermelo-Fraenkel (ZF) sets as abstract individuals, but also identify such mathematical relations as the *successor* relation in Peano Arithmetic, the *greater than* relation in real number theory, and the *membership* relation of ZF, as *abstract relations* (i.e., relations that encode properties of relations). To do this, we need a theory of abstract relations that can encode properties of relations. Such a theory can be obtained by reformulating the language, logic, and proper axioms of the present system within relational type theory.

Using typed object theory, we shall analyze a mathematical theory as a certain kind of situation. As such, mathematical theories become identified as abstract objects that encode only propositional properties. We shall then define *truth in a mathematical theory* T in the same way that we defined truth in a situation, truth at a world, and truth in a story, namely, a proposition p is true in theory T just in case T encodes the propositional property $[\lambda x\, p]$. Our analysis will preserve the pretheoretic constraint on mathematical theories, namely, that propositions logically implied by propositions true in the theory are also true in the theory. Given this analysis, we shall then go on to analyze the ob-

[387]See Chapter 10, item (309), where we distinguished theoretical mathematical objects from natural mathematical objects.

jects and relations of *arbitrary*, mathematical theories.

To make all this possible, we work our way to the formulation of the following comprehension principle for abstract objects, which is expressible in a language built with respect to the simple theory of relational types t:

> $\exists x^t(A!^{\langle t\rangle}x \,\&\, \forall F^{\langle t\rangle}(xF \equiv \varphi))$, where t is any type and φ is any formula having no free occurrences of the variable x of type t.

Then we shall not only identify (theoretical) mathematical objects as individuals of type i, but also identify (theoretical) unary and binary mathematical relations as abstract objects of type $\langle i\rangle$ and type $\langle i,i\rangle$, respectively. For example, the number π of real number theory will be identified as that abstract individual that encodes just the properties exemplified by π in real number theory. Similarly, the relation $\in$ of Zermelo-Fraenkel set theory will be identified as that abstract relation that encodes just the properties of relations exemplified by the relation $\in$ in ZF. Those readers familiar with relational type theory may want to glance at the notation used here for relational types and then skip ahead to Section (15.6), where we begin the exposition of our analysis of mathematical theories, objects, and relations.

15.1 The Language and Its Interpretation

(**924**) **Metadefinition:** Types. We begin with a definition of the *types*. These categorize both the terms of our language and the entities that such terms may denote.[388]

Let i be the sole *underived* or *primitive* type for *individuals*, where individuals are intuitively understood to be non-predicable entities. Then, using $t, t_1, t_2, \ldots$ as variables ranging over types, we may define *type* as follows:[389]

[388] We are here categorizing the types of *simple relational type theory*. Though Russell developed *ramified* type theory in 1906 and 1908, Church (1974, 21) credits the development of simple type theory to Chwistek 1921 and 1922, Ramsey 1926, and Carnap 1929. However, the notation for simple type theory seems to have stabilized in Orey 1959 (73), though the 'derived' (i.e., complex) types are defined only for $n \geq 1$. Church 1974 (25) and Gallin 1975 (68) use notation similar to Orey 1959, but let $n = 0$ to obtain the derived type for propositions. This notation was used in Zalta 1988 (Appendix). The relational types formulated in Zalta 1983 (Chapter VI) were essentially the same as those in Zalta 1988, but the primitive type p used in 1983 was abandoned in 1988 when I realized it could be defined as the empty derived type.

It may also be worth noting in Orey 1959 (73), Church 1974 (26), Gallin 1975 (72), and Muskens 1989a, the entities in the domains of the relational types were taken to be sets of n-tuples or functions. However, in Zalta 1983, 1988, and 2020, they were taken to be primitive, *intensional* relations. See Zalta 2020 for a fuller discussion of the history of interpretations of relational type theory.

[389] These types were used in Zalta 1988, Chapters 9, 12, and the Appendix. That work was a simpler variant of the type theory used in Zalta 1983, Chapter V. See also the derived types for the system ML_p in Gallin 1975, 68.

(.1) 'i' is a type.

(.2) If $t_1, \ldots, t_n$ are any types ($n \geq 0$), $\ulcorner\langle t_1, \ldots, t_n\rangle\urcorner$ is a type.

The *derived* or *complex* types of the form $\langle t_1, \ldots, t_n\rangle$ in (.2) are the types of *n-place relations* among objects of type $t_1, \ldots, t_n$, in that order. A derived type of the form $\langle t\rangle$ is the type for *properties* of objects of type t. The empty, derived type $\langle\,\rangle$ is the type for *propositions* or *states of affairs*.

For most of this chapter, we shall be focusing on ordinary and abstract objects of the following types:

i	individuals
$\langle\,\rangle$	propositions = states of affairs
$\langle i\rangle$	properties of individuals
$\langle i,i\rangle$	binary relations among individuals
$\langle i,\langle\,\rangle\rangle$	binary relations between individuals and propositions
$\langle i,i,i\rangle$	tertiary relations among individuals
$\langle\langle i\rangle\rangle$	properties of properties of individuals
$\langle\langle i,i\rangle\rangle$	properties of binary relations among individuals

Thus, the type of propositional attitude verbs such as *belief* already appear in the list; these are relations of type $\langle i,\langle\,\rangle\rangle$, i.e., a binary relation between an individual and a proposition or state of affairs. But other natural language expressions can be typed using the above scheme. For example, a sentential adverb might be assigned type $\langle\langle\,\rangle,\langle\,\rangle\rangle$, i.e., as a binary relation between propositions, while a predicate adverb might be assigned the type $\langle\langle i\rangle,\langle i\rangle\rangle$, i.e., a binary relation between properties of individuals.

It should be noted that for most of this chapter, we shall not distinguish propositions and states of affairs. For most purposes, it makes no difference whether we call 0-ary relations 'propositions' or 'states of affairs'. However, when it comes time to analyze *propositional attitudes* and *propositional attitude reports*, it may prove useful to regard 0-ary relations of type $\langle\,\rangle$ in one of two ways, depending on whether we are discussing *de re* or *de dicto* interpretations of attitude reports. When discussing *de re* reports and our interest lies in the *truth of a belief* or the *satisfaction of a desire*, it is useful to call the 0-ary relation upon which truth or satisfaction depends a 'state of affairs'. When discussing *de dicto* reports and our interest lies in the mental representation of a state of affairs, it useful to call the 0-ary relation that serves to *represent* the state of affairs believed or desired a 'proposition'. In both cases, we're still referencing 0-ary relations, but this usage will tie propositions more closely to the content of mental state representations and tie states of affairs more closely to the state of the world being represented.

(**925**) **Metadefinition:** Simple Terms of Type t. We shall assume that for every type t, there are two kinds of *simple* terms of that type, namely, *constants* and *variables*:

Constants of type t: $a_1^t, a_2^t, \ldots$

Variables of type t: $x_1^t, x_2^t, \ldots$

Consequently, there is a denumerable list of constants and variables, with each symbol indexed both by type and its place in the enumeration. The resulting language, however, becomes difficult to read, and so we shall almost always avail ourselves of other symbols to facilitate understanding. In particular, we shall continue to write the relation symbol in a predication in upper case. And, we frequently indicate, in the text, the type of the symbols used in a formula about to be displayed, so that we may display the formula without the distraction of the type superscripts.

Note that in the second-order fragment studied in Chapters 7 – 14, the language included terms for individuals and terms for n-ary relations among individuals ($n \geq 0$). However, it included no terms for higher-order relations. Thus, all of the relations studied in those chapters were of type $\langle i, \ldots, i\rangle$, in which there are n occurrences of i, for some $n \geq 0$.

(**926**) **Metadefinitions:** Syncategorematic Expressions for Typed Object Theory. A syncategorematic expression is an expression that is neither a term (i.e., assigned a denotation) nor a formula (i.e., assigned truth conditions), but which nevertheless represents a primitive notion. We use following syncategorematic expressions in both the type-theoretic language and the second-order fragment, where α^t ranges over variables of type t:

(.1) Unary Formula-Forming Operators:
$\neg$ ('it is not the case that' or 'it is false that')
$\Box$ ('necessarily' or 'it is necessary that')
$\mathcal{A}$ ('actually' or 'it is actually the case that')

(.2) Binary Formula-Forming Operator:
$\rightarrow$ ('if ..., then ...')

(.3) Variable-Binding Formula-Forming Operator:
$\forall\alpha^t$ ('every α^t is such that')
for every variable α^t of any type t

(.4) Variable-Binding Individual-Term-Forming Operator:
$\imath\alpha^t$ ('the α^t such that')
for every variable α^t of any type t

(.5) Variable-Binding n-ary Relation-Term-Forming Operators ($n \geq 0$):
$\lambda\alpha^{t_1}\ldots\alpha^{t_n}$ ('being $\alpha^{t_1}\ldots\alpha^{t_n}$ such that', or when $n = 0$, 'that')
for any distinct variables $\alpha^{t_1},\ldots,\alpha^{t_n}$ with types $t_1,\ldots,t_n$, respectively

These primitive, syncategorematic expressions are referenced in the definition of the syntax of our language and are used to define complex formulas and complex terms. In what follows, we sometimes call $\neg$ the *negation* operator, $\mathcal{A}$ the *actuality* operator, $\Box$ the *necessity* operator, $\to$ the *conditional*; $\forall\alpha^t$ a *universal quantifier*, $\iota\alpha^t$ the *definite description* operator; and $\lambda\alpha^{t_1}\ldots\alpha^{t_n}$ the *relation abstraction* operator. By convention, in any conditional formula of the form $\varphi \to \psi$, we say φ is the antecedent, and ψ the consequent, of the conditional.

(**927**) **Definition:** The Language of Typed Object Theory. In addition using α^t as a metavariable ranging over variables of type t, we use the following Greek metavariables:

- τ^t ranges over *terms of type t*, where t is any type
- φ ranges over *formulas*
- $\Pi^{\langle t_1,\ldots,t_n\rangle}$ is sometimes used as a more easily readable substitute for $\tau^{\langle t_1,\ldots,t_n\rangle}$ ($n \geq 0$)

We then define, by simultaneous recursion, the *terms of type t* and *formulas* of our type-theoretic language as follows:

Base Clauses:

- Terms: Simple terms (i.e., constants and variables) of type t are terms of type t, for every type t.
- Terms: $E!^t$ is a distinguished unary relation constant, for every type $t \neq i$.
- Formulas: Where Π is a constant, a variable, or a description of type $\langle\,\rangle$, Π is a formula.
- Formulas: Where $\Pi^{\langle t_1,\ldots,t_n\rangle}$ is a term of type $\langle t_1,\ldots,t_n\rangle$, and $\tau^{t_1},\ldots,\tau^{t_n}$ are terms of types $t_1,\ldots,t_n$, respectively, for $n \geq 1$, then an expression of the form $\Pi\tau^{t_1}\ldots\tau^{t_n}$ is an *exemplification* formula.
- Formulas: Where $\tau^{t_1},\ldots,\tau^{t_n}$ are terms of types $t_1,\ldots,t_n$, respectively, and $\Pi^{\langle t_1,\ldots,t_n\rangle}$ is a term of type $\langle t_1,\ldots,t_n\rangle$, for $n \geq 1$, then an expression of the form $\tau^{t_1}\ldots\tau^{t_n}\Pi$ is an *encoding* formula.

Recursive Clauses:

- Formulas: If φ and ψ are any formulas, and α^t is any variable of type t, then $[\lambda\,\varphi]$, $\neg\varphi$, $\varphi \to \psi$, $\forall\alpha^t\varphi$, $\Box\varphi$, and $\mathcal{A}\varphi$ are formulas.

- Terms: Where φ is any formula and α^t is any variable of type t, then $\imath\alpha^t\varphi$ is a term of type t.

- Terms: Where φ is any formula and $\alpha^{t_1},\ldots,\alpha^{t_n}$ ($n \geq 1$) are any *distinct* variables having types $t_1,\ldots,t_n$, respectively, then:

  (a) $[\lambda\alpha^{t_1}\ldots\alpha^{t_n}\,\varphi]$ is a term of type $\langle t_1,\ldots,t_n\rangle$, and

  (b) φ itself is a term of type $\langle\,\rangle$.

We say that an expression is a *term* just in case it is a term of type t, for some type t.

As noted previously, we will make it easier to read formulas of the language by using informal notation. In particular, we use the upper case, italic letters $P,Q,\ldots$ and $F,G,\ldots$ to abbreviate constants and variables, respectively, of some antecedently-specified relational type of the form $\langle t_1,\ldots,t_n\rangle$, for $n \geq 1$. We use the italic letters $p,q,\ldots$ as variables ranging over propositions (i.e., instead of the variables $x_1^{\langle\rangle},x_2^{\langle\rangle},\ldots$).

So, for example, if we suppose that F is a variable having a type of the form $\langle t_1,t_2\rangle$, and suppose that a is a constant of type t_1 and x is a constant of type t_2, then Fax is a well-formed exemplification formula and axF is a well-formed encoding formula. Similarly, if B is a constant of type $\langle i,\langle\,\rangle\rangle$ that represents the binary relation of *belief*, a is a constant of type i, and p is a variable for a proposition, then Bap asserts that *a believes p* (or *a believes* the proposition p), and $Ba[\lambda\,p]$ asserts that *a believes that p*.

(**928**) **Metadefinition:** A BNF for the Type-Theoretic Language. In order to present a BNF definition for the language of typed object theory, we make use of the following metavariables:

δ^t	primitive constants of type t
α^t	variables of type t
τ^t	terms of type t
φ	formulas

If we now temporarily let these Greek metavariables serve as names of the syntactic categories of the expressions over which they range, then we may succinctly state the context-free grammar of our type-theoretic language using

Backus-Naur Form (BNF), as follows:

$$
\begin{array}{rrcl}
 & \delta^t & ::= & a_1^t, a_2^t, \ldots \quad (E!^{\langle t \rangle} \text{ a distinguished constant, for every } t) \\
 & \alpha^t & ::= & x_1^t, x_2^t, \ldots \\
 & Base^t & ::= & \delta^t \mid \alpha^t \mid \imath\alpha^t\varphi \\
 & \tau^i & ::= & Base^i \\
(n \geq 1) & \tau^{\langle t_1,\ldots,t_n \rangle} & ::= & Base^{\langle t_1,\ldots,t_n \rangle} \mid [\lambda\alpha^{t_1} \ldots \alpha^{t_n}\,\varphi] \quad (\alpha^{t_1} \ldots \alpha^{t_n} \text{ pairwise distinct}) \\
 & \varphi & ::= & Base^{\langle\rangle} \mid \tau^{\langle t_1,\ldots,t_n \rangle}\tau^{t_1} \ldots \tau^{t_n}\ (n \geq 1) \mid \tau^{t_1} \ldots \tau^{t_n}\tau^{\langle t_1,\ldots,t_n \rangle}\ (n \geq 1) \mid \\
 & & & [\lambda\,\varphi] \mid (\neg\varphi) \mid (\varphi \to \varphi) \mid \forall\alpha^t\varphi \mid (\Box\varphi) \mid (\mathcal{A}\varphi) \\
 & \tau^{\langle\rangle} & ::= & \varphi
\end{array}
$$

It is insightful to recognize that one could, in the above BNF, replace φ everywhere by $\tau^{\langle\rangle}$, except on the last line, which would then be reversed to read $\varphi ::= \tau^{\langle\rangle}$. This alternative BNF would first introduce the terms of every type and then define the formulas as terms of type $\langle\rangle$. However, it is a considered choice, based on ease of readability and understanding, to write the BNF as the above.

Given this BNF, we may again say that a term τ is any expression of type τ^t, for some type t. As mentioned previously, we often use Π instead of τ as a metavariable ranging over terms of type t when $t \neq i$. Such terms are said to be of *relational type*, where these are to be distinguished from terms for individuals having type i.

It may not be immediately apparent that the language defined in (4) is a special case of the above language. To see this how the above reduces to (4) as a special case, restrict the definition of *types* to the type i for individuals and the types of the form $\langle i,\ldots,i \rangle$ for n-ary relations among individuals, for any n-long string of is ($n \geq 0$). So we may eliminate from the above BNF every constant, variable, and complex term of every other type, though we can make the constants and variables for individuals and relations among individuals introduced in (4) explicit. Then it can be seen that (a) the constants and variables of (4) are restricted to those of types i and $\langle i,\ldots,i \rangle$, (b) the only definite descriptions in (4) are those of type i (in which the $\imath$-operator binds a single variable of type i), and (c) the only λ-expressions in (4) are those of type $\langle i,\ldots,i \rangle$ (in which the λ-operator binds pairwise distinct variables of types i). Thus, the lines for formulas φ and 0-ary relation terms $\tau^{\langle\rangle}$ in the above BNF reduce to the lines for φ and Π^0, respectively, in (4).

(**929**) **Remark:** A Semantic Interpretation. Though the language of object theory was designed to be transparent, it would serve well to specify, in complete generality, the denotations of the terms and the truth conditions of the formulas. Again, we assume as a metalanguage, the first-order language of set theory with urelements, supplemented by a metalinguistic $\bar{\epsilon}$ operator of the kind described in the opening paragraph of Section 5.2.1 and in footnote 37. This metalanguage allows us to describe a general structure and define a formal semantic interpretation of the language, including general semantic definitions of truth, validity, and logical consequence. This will give the reader an independent picture that provides a representation or model of *what the axioms assert* (assuming them to be true) from the standpoint of set theory + urelements; it does *not* constitute a model that demonstrates that the axioms are jointly true (we leave the construction of a model to some other occasion – see for example, the initial efforts in Leitgeb, Nodelman, and Zalta m.s.). Rather one should think of the axioms as placing constraints that any structure of the kind described below must satisfy.

Consider the following structure:

$$\mathcal{I} = \langle \mathbf{D}, \mathbf{W}, \boldsymbol{T}, \boldsymbol{F}, \mathbf{ext}_{\boldsymbol{w}}, \mathbf{enc}_{\boldsymbol{w}}, \mathbf{ex}_{\boldsymbol{w}}, \mathbf{V}, \mathbf{C} \rangle,$$

where:

- $\mathbf{D}$ is the general union of non-empty domains $\mathbf{D}_t$, for every type t; i.e., $\mathbf{D} = \bigcup_t \mathbf{D}_t$. We often use $\boldsymbol{o}^t$ as a variable ranging over the elements of $\mathbf{D}_t$; use $\boldsymbol{r}$ as a variable ranging over the elements of $\mathbf{D}_{\langle t_1,\ldots,t_n \rangle}$, where $t_1, \ldots, t_n$ are any types and $n \geq 1$; and use $\boldsymbol{p}$ as a variable ranging over the elements of $\mathbf{D}_{\langle \rangle}$,

- $\mathbf{W}$ is a non-empty set of possible worlds with a *distinguished element* $\boldsymbol{w}_0$; we use $\boldsymbol{w}$ as a variable ranging over the elements of $\mathbf{W}$,

- $\boldsymbol{T}$ is the truth-value The True,

- $\boldsymbol{F}$ is the truth-value The False,

- $\mathbf{ext}_{\boldsymbol{w}}$ is a binary *exemplification extension* function indexed to its second argument; $\mathbf{ext}_{\boldsymbol{w}}$ maps each relation $\boldsymbol{r}$ in $\mathbf{D}_{\langle t_1,\ldots,t_n \rangle}$ ($n \geq 1$) and world $\boldsymbol{w}$ to a set of n-tuples whose elements have types $t_1, \ldots, t_n$, respectively, so that $\mathbf{ext}_{\boldsymbol{w}}(\boldsymbol{r})$ serves as the exemplification extension of $\boldsymbol{r}$ at $\boldsymbol{w}$,[390]

- $\mathbf{enc}_{\boldsymbol{w}}$ is a binary *encoding extension* function indexed to its second argument; $\mathbf{enc}_{\boldsymbol{w}}$ maps each relation $\boldsymbol{r}$ in $\mathbf{D}_{\langle t_1,\ldots,t_n \rangle}$ ($n \geq 1$) and world $\boldsymbol{w}$ to a set of n-tuples whose elements have types $t_1, \ldots, t_n$, respectively, so that $\mathbf{enc}_{\boldsymbol{w}}(\boldsymbol{r})$ serves as the encoding extension of $\boldsymbol{r}$ at $\boldsymbol{w}$,

- $\mathbf{ex}_{\boldsymbol{w}}$ is a binary *extension* function indexed to its second argument; $\mathbf{ex}_{\boldsymbol{w}}$ maps each proposition $\boldsymbol{p}$ in $\mathbf{D}_{\langle \rangle}$ and world $\boldsymbol{w}$ to one of the truth-values ($\boldsymbol{T}$ or $\boldsymbol{F}$) so that $\mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{p})$ serves as the extension of $\boldsymbol{p}$ at $\boldsymbol{w}$,

[390] By convention, $\mathbf{ext}_{\boldsymbol{w}}$ maps each relation unary relation $\boldsymbol{r}$ in $\mathbf{D}_{\langle t \rangle}$ ($n \geq 1$) and world $\boldsymbol{w}$ to a subset of $\mathbf{D}_t$.

- $\mathbf{V}$ is an interpretation function that assigns each the primitive constant of type t to an element of the domain $\mathbf{D}_t$, and

- $\mathbf{C}$ is a choice function that takes, as argument, any semantic formula $\boldsymbol{A}$ having a single free variable that ranges over some domain $\mathbf{D}_t$, for $t \neq i$, and returns an arbitrary but determinate value in $\mathbf{D}_t$ that satisfies $\boldsymbol{A}$ if there is one, and is undefined otherwise. Then where $\bar{\varepsilon}\boldsymbol{r}^n\boldsymbol{A}$ is any semantic $\bar{\varepsilon}$-term in which $\boldsymbol{r}^n$ is a semantic variable that ranges over the n-ary relations ($n \geq 0$) in the domain $\mathbf{D}_{\langle t_1,\ldots,t_n\rangle}$, the object $\mathbf{C}(\boldsymbol{A})$ is an arbitrarily chosen entity of type $\langle t_1,\ldots,t_n\rangle$ that satisfies $\boldsymbol{A}$, if there is one, which serves as the value of the term. For example, if $\boldsymbol{A}$ has $\boldsymbol{r}$ free and $\boldsymbol{r}$ ranges over relations in $\mathbf{D}_{\langle i,i\rangle}$ (i.e., ranges over binary relations among individuals), then the semantic term $\bar{\varepsilon}\boldsymbol{r}^n\boldsymbol{A}$ denotes $\mathbf{C}(\boldsymbol{A})$, where the latter is an arbitrary but determinate relation in $\mathbf{D}_{\langle i,i\rangle}$ that satisfies $\boldsymbol{A}$, if there is one. Similarly, if $\boldsymbol{A}$ has $\boldsymbol{p}$ free, where $\boldsymbol{p}$ ranges over $\mathbf{D}_{\langle\rangle}$, then $\bar{\varepsilon}\boldsymbol{p}\boldsymbol{A}$ denotes $\mathbf{C}(\boldsymbol{A})$, where the latter is an arbitrary but determinate proposition in $\mathbf{D}_{\langle\rangle}$ that satisfies $\boldsymbol{A}$, if there is one.

Given such a structure $\mathcal{I}$, let $\boldsymbol{w}$ range over the primitive possible worlds in $\mathbf{W}$, and let f be a *assignment function* relative to $\mathcal{I}$ that assigns to each variable α^t an element of the domain $\mathbf{D}_t$. (For ease of readability, we always omit the index on f that relativizes it to $\mathcal{I}$.) Then we shall assign denotations to the terms and truth conditions to the formulas by defining the following notions simultaneously:

$\boldsymbol{d}_{\mathcal{I},f}(\tau)$, i.e., the denotation of τ relative to $\mathcal{I}$ and f

$\boldsymbol{w} \models_{\mathcal{I},f} \varphi$, i.e., *under* $\mathcal{I}$ and f, φ *is true at* $\boldsymbol{w}$

The definitions are given in full below but note that, in what follows, we are re-purposing the symbol $\models$ for the semantics. When we use $\models$ in a semantic context in what follows, it is to be understood as representing a semantic notion, and not the object-theoretic notion p is true in s ($s \models p$) defined in (470).

Intuitively, $\boldsymbol{d}_{\mathcal{I},f}$ is a partial denotation function which, relative to an interpretation $\mathcal{I}$ and variable assignment f, assigns to every term τ of type t an element of the domain $\mathbf{D}_t$ if τ is significant, and nothing otherwise. And, $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ states the truth conditions of φ at world $\boldsymbol{w}$, relative to $\mathcal{I}$ and f. Now let:

- $\mathcal{I}$ be any interpretation and f be any assignment function,

- $\mathbf{V}$ be the interpretation function of $\mathcal{I}$,

- $f[\alpha^t/\boldsymbol{o}^t]$ be the variable assignment just like f except that it assigns the entity $\boldsymbol{o}^t$ to the variable α^t,[391] and
- $f[\alpha^{t_i}/\boldsymbol{o}^{t_i}]_{i=1}^n$ be the variable assignment just like f but which assigns the entities $\boldsymbol{o}^{t_1},\ldots,\boldsymbol{o}^{t_n}$, respectively, to the variables $\alpha^{t_1},\ldots,\alpha^{t_n}$, for $1 \leq i \leq n$

And let us adopt the convention of omitting the type index on a symbol after its first use in a semantic formula whenever it can be done without ambiguity. Then the simultaneous definition of denotation and world-relative truth, relative to $\mathcal{I}$ and f, proceeds as follows:

Base Clauses

D1. If τ is a constant of type t, then $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \mathbf{V}(\tau)$

D2. If τ is a variable of type t, then $\boldsymbol{d}_{\mathcal{I},f}(\tau) = f(\tau)$

T1. If φ is a formula in $Base^{\langle\rangle}$, i.e., if φ is a constant, variable, or description of type $\langle\,\rangle$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if $\exists \boldsymbol{p}^{\langle\rangle}(\boldsymbol{p} = \boldsymbol{d}_{\mathcal{I},f}(\varphi) \;\&\; \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{p}) = \boldsymbol{T})$

T2. If φ is a formula of the form $\Pi^{\langle t_1,\ldots,t_n\rangle}\tau^{t_1}\ldots\tau^{t_n}$ $(n \geq 1)$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if $\exists \boldsymbol{r}^{\langle t_1,\ldots,t_n\rangle}\exists \boldsymbol{o}^{t_1}\ldots\exists \boldsymbol{o}^{t_n}(\boldsymbol{r} = \boldsymbol{d}_{\mathcal{I},f}(\Pi) \;\&\; \boldsymbol{o}^{t_1} = \boldsymbol{d}_{\mathcal{I},f}(\tau^{t_1}) \;\&\; \ldots \;\&\; \boldsymbol{o}^{t_n} = \boldsymbol{d}_{\mathcal{I},f}(\tau^{t_n}) \;\&\; \langle \boldsymbol{o}^{t_1},\ldots,\boldsymbol{o}^{t_n}\rangle \in \mathbf{ext}_{\boldsymbol{w}}(\boldsymbol{r}))$

T3. If φ is a formula of the form $\tau^{t_1}\ldots\tau^{t_n}\Pi^{\langle t_1,\ldots,t_n\rangle}$ $(n \geq 1)$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if $\exists \boldsymbol{o}^{t_1}\ldots\exists \boldsymbol{o}^{t_n}\exists \boldsymbol{r}^{\langle t_1,\ldots,t_n\rangle}(\boldsymbol{o}^{t_1} = \boldsymbol{d}_{\mathcal{I},f}(\tau^{t_1}) \;\&\; \ldots \;\&\; \boldsymbol{o}^{t_n} = \boldsymbol{d}_{\mathcal{I},f}(\tau^{t_n}) \;\&\; \boldsymbol{r} = \boldsymbol{d}_{\mathcal{I},f}(\Pi) \;\&\; \langle \boldsymbol{o}^{t_1},\ldots,\boldsymbol{o}^{t_n}\rangle \in \mathbf{enc}_{\boldsymbol{w}}(\boldsymbol{r}))$

Recursive Clauses

T4. If φ is a formula of the form $[\lambda\, \psi]$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if $\boldsymbol{w} \models_{\mathcal{I},f} \psi$

T5. If φ is a formula of the form $\neg\psi$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if it is not the case that $\boldsymbol{w} \models_{\mathcal{I},f} \psi$, i.e., iff $\boldsymbol{w} \not\models_{\mathcal{I},f} \psi$

T6. If φ is a formula of the form $\psi \rightarrow \chi$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if either it is not the case that $\boldsymbol{w} \models_{\mathcal{I},f} \psi$ or it is the case that $\boldsymbol{w} \models_{\mathcal{I},f} \chi$, i.e., iff either $\boldsymbol{w} \not\models_{\mathcal{I},f} \psi$ or $\boldsymbol{w} \models_{\mathcal{I},f} \chi$

[391] This can be defined formally in one of two ways, suppressing the type index. If an assignment function f is represented as a set of ordered pairs, then where α is a variable and $\boldsymbol{o}$ is an entity from the domain over which α ranges:

$$f[\alpha/\boldsymbol{o}] = (f \sim \langle \alpha, f(\alpha)\rangle) \cup \{\langle \alpha, \boldsymbol{o}\rangle\}$$

I.e., $f[\alpha/\boldsymbol{o}]$ is the result of removing the pair $\langle \alpha, f(\alpha)\rangle$ from f and replacing it with the pair $\langle \alpha, \boldsymbol{o}\rangle$.

Alternatively, we can define $f[\alpha/\boldsymbol{o}]$ functionally, where β is a variable ranging over the same domain as α, as:

$$f[\alpha/\boldsymbol{o}](\beta) = \begin{cases} f(\beta), & \text{if} \beta \neq \alpha \\ \boldsymbol{o}, & \text{if} \beta = \alpha \end{cases}$$

T7. If φ is a formula of the form $\forall\alpha^t\psi$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if $\forall\boldsymbol{o}^t(\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{o}]} \psi)$

T8. If φ is a formula of the form $\Box\psi$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if $\forall\boldsymbol{w}'(\boldsymbol{w}' \models_{\mathcal{I},f} \psi)$

T9. If φ is a formula of the form $\mathcal{A}\psi$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if $\boldsymbol{w}_0 \models_{\mathcal{I},f} \psi$.

D3. If τ is a description of the form $\imath\alpha^t\varphi$, then

$$\boldsymbol{d}_{\mathcal{I},f}(\tau) = \begin{cases} \boldsymbol{o}^t, \text{ if } \boldsymbol{w}_0 \models_{\mathcal{I},f[\alpha/\boldsymbol{o}]} \varphi \;\&\; \forall\boldsymbol{o}'(\boldsymbol{w}_0 \models_{\mathcal{I},f[\alpha/\boldsymbol{o}']} \varphi \rightarrow \boldsymbol{o}' = \boldsymbol{o}) \\ \text{undefined, otherwise} \end{cases}$$

where $\boldsymbol{o}'$ also ranges over the entities in $\mathbf{D}_t$

D4. If τ is an n-ary λ-expression ($n \geq 1$) of the form $[\lambda\alpha^{t_1} \ldots \alpha^{t_n}\, \varphi]$, then

$$\boldsymbol{d}_{\mathcal{I},f}(\tau) = \begin{cases} \bar{\varepsilon}\boldsymbol{r}^{\langle t_1,\ldots,t_n\rangle}\forall\boldsymbol{w}\forall\boldsymbol{o}^{t_1}\ldots\forall\boldsymbol{o}^{t_n}(\langle\boldsymbol{o}^{t_1},\ldots,\boldsymbol{o}^{t_n}\rangle \in \mathbf{ext}_{\boldsymbol{w}}(\boldsymbol{r}) \;\equiv\; \boldsymbol{w} \models_{\mathcal{I},f[\alpha^{t_i}/\boldsymbol{o}^{t_i}]_{i=1}^n} \varphi), \\ \quad\text{if there is one} \\ \text{undefined, otherwise} \end{cases}$$

where $\bar{\varepsilon}\boldsymbol{r}\boldsymbol{A} = \mathbf{C}(\boldsymbol{A})$ and $\mathbf{C}$ is the choice function of the interpretation.

D5. If τ is an 0-ary λ-expression of the form $[\lambda\, \varphi]$, then

$$\boldsymbol{d}_{\mathcal{I},f}(\tau) = \bar{\varepsilon}\boldsymbol{p}^{\langle\rangle}\forall\boldsymbol{w}(\mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{p}) = \boldsymbol{T} \;\equiv\; \boldsymbol{w} \models_{\mathcal{I},f} \varphi)$$

where $\bar{\varepsilon}\boldsymbol{p}\boldsymbol{A} = \mathbf{C}(\boldsymbol{A})$ and $\mathbf{C}$ is the choice function of the interpretation.

D6. If τ is a term of type $\langle\,\rangle$, i.e., if τ is a formula φ, then:

- if φ is a formula in $Base^{\langle\rangle}$, then $\boldsymbol{d}_{\mathcal{I},f}(\tau)$ is given by D1 – D3
- if φ is a formula of the form $[\lambda\, \varphi]$, then $\boldsymbol{d}_{\mathcal{I},f}(\tau)$ is given by D5
- if φ is a formula of any other form, then $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},f}([\lambda\, \varphi])$

Now where $\mathcal{I}$ and f are given and $\boldsymbol{w}_0$ is the distinguished actual world of the domain of possible worlds $\mathbf{W}$ in $\mathcal{I}$, we say that φ is *true under* $\mathcal{I}$ and f ('true$_{\mathcal{I},f}$') if and only if under $\mathcal{I}$ and f, φ is *true at* $\boldsymbol{w}_0$. That is, using the formal notation $\models_{\mathcal{I},f} \varphi$ for the definiendum, we have:

$\models_{\mathcal{I},f} \varphi$ if and only if $\boldsymbol{w}_0 \models_{\mathcal{I},f} \varphi$

And we now say that φ is *true under* $\mathcal{I}$ just in case for every f, φ is true under $\mathcal{I}$ and f:

$\models_{\mathcal{I}} \varphi \;=_{df}\; \forall f(\models_{\mathcal{I},f} \varphi)$

Thus, if φ is not true under $\mathcal{I}$, then some assignment f is such that $\boldsymbol{w}_0 \not\models_{\mathcal{I},f} \varphi$ and we write $\not\models_{\mathcal{I}} \varphi$. We say that a formula φ is *false under* $\mathcal{I}$ if and only if no assignment function f is such that $\models_{\mathcal{I},f} \varphi$, i.e., iff no assignment function f is such that $\boldsymbol{w}_0 \models_{\mathcal{I},f} \varphi$. So open formulas may be neither true under $\mathcal{I}$ nor false under $\mathcal{I}$, whereas a sentence (i.e., a closed formula) will be either true under $\mathcal{I}$ or false under $\mathcal{I}$.

In the usual manner, we say that φ is *valid* or *logically true* if and only if φ is true under every interpretation $\mathcal{I}$, i.e.,

$$\models \varphi \ =_{df} \ \forall \mathcal{I}(\models_{\mathcal{I}} \varphi)$$

Clearly, given our previous definitions, it follows that:

$\models \varphi$ if and only if for every $\mathcal{I}$ and f, $\models_{\mathcal{I},f} \varphi$, i.e.,

$\models \varphi$ if and only if for every $\mathcal{I}$ and f, $\boldsymbol{w}_0 \models_{\mathcal{I},f} \varphi$

In what follows, when we say that a schema is valid, we mean that all of its instances are valid. Clearly, if a formula φ is not valid, then for some interpretation $\mathcal{I}$ and assignment f, $\boldsymbol{w}_0 \not\models_{\mathcal{I},f} \varphi$.

Finally, we conclude the definitions for a general interpretation with several more traditional definitions:

- φ is *satisfiable* if and only if there is some interpretation $\mathcal{I}$ and assignment f such that φ is $\text{true}_{\mathcal{I},f}$, i.e., iff $\exists \mathcal{I} \exists f(\models_{\mathcal{I},f} \varphi)$.
- φ *logically implies* ψ (or ψ is a *logical consequence* of φ) just in case, for every interpretation $\mathcal{I}$ and assignment f, if φ is $\text{true}_{\mathcal{I},f}$, then ψ is $\text{true}_{\mathcal{I},f}$:

  $$\varphi \models \psi \ =_{df} \ \forall \mathcal{I} \forall f(\models_{\mathcal{I},f} \varphi \rightarrow \ \models_{\mathcal{I},f} \psi)$$

- φ and ψ are *logically equivalent* just in case both $\varphi \models \psi$ and $\psi \models \varphi$:

  $$\varphi \mathrel{=\!\!\mid\!\models} \psi =_{df} \ \varphi \models \psi \ \& \ \psi \models \varphi$$

- φ is a *logical consequence* of a set of formulas Γ just in case, for every interpretation $\mathcal{I}$ and assignment f, if every member of Γ is $\text{true}_{\mathcal{I},f}$, then φ is $\text{true}_{\mathcal{I},f}$:

  $$\Gamma \models \varphi \ =_{df} \ \forall \mathcal{I} \forall f[\forall \psi(\psi \in \Gamma \rightarrow \ \models_{\mathcal{I},f} \psi) \rightarrow \ \models_{\mathcal{I},f} \varphi]$$

(**930**) **Metadefinitions:** Abstract Syntax. In what follows, we use the notational conventions developed in (5), suitably adapted to the type-theoretic context. Moreover, the metatheoretic notions of abstract syntax that were defined with respect to second-order object theory can all be straightforwardly adapted to the more general type-theoretic framework. We leave this as an exercise, though on occasion we may note where these notions have to be adjusted slightly if they are to be used in the more general type-theoretic language. In what follows, therefore, we make use of the following notions of abstract syntax:

- *subformula* (6)
- *subterms* and *primary terms* (7)
- *operator scope* and *free/bound* occurrences of variables (8)
- *encoding position* (9.1)
- *core* λ-expressions (9.2)
- *open/closed* formulas/terms (10)
- *closures* (11)

We'll henceforth reference these metadefinitions in what follows by both the present and original item number. So, for example, the definition of subformula will be referenced as (930) [6].

> **Observation:** Consider how one would adapt the definition of a *core* λ expression:
>
> > If $t_1,\ldots,t_n$ are any types, and $\alpha_1,\ldots,\alpha_n$ are any variables with types $t_1,\ldots t_n$, respectively, then $[\lambda\alpha_1\ldots\alpha_n\varphi]$ is a *core* λ-expression if and only if no variable bound by the λ occurs in encoding position in φ.
>
> Of interest here is the fact that in typed object theory, the λ can bind variables of any type, whereas in second-order object theory, the λ may bind only individual variables. This adapted definition of a core λ-expression therefore gives rise to new cases of λ-expressions that *fail* to be core λ-expressions, as well as new cases that satisfy the definition. For example, let x be a variable of some type t and F be a variable of type $\langle t\rangle$. Then not only does $[\lambda x\, xF]$ fail to be a core λ-expression, but so does $[\lambda F\, xF]$. In each case, the λ binds a variable that occurs in encoding position in the matrix. Note, however, that when $O!$ has type $\langle t\rangle$ and x and F have the same type as in the previous example, then both $[\lambda F\, x[\lambda z\, Fz]]$ and $[\lambda F\, O!F\, \&\, x[\lambda z\, Fz]]$ are core λ-expressions. The variable F, in each case, doesn't occur in encoding position in the matrix. Our derivation system will *not* allow one to substitute F for $[\lambda z\, Fz]$ unless it is known that $O!F$, since η-Conversion will assert the identity $[\lambda z\, Fz] = F$ only under the condition that F is an ordinary relation; see the statement of η-Conversion in (935.27). So we won't be able to derive the identity

$[\lambda F\, x[\lambda z\, Fz]] = [\lambda F\, xF]$. The term on the left of this identity isn't a core λ-expression and won't denote, whereas the term on the right of this identity is a core λ-expression and will denote. But we will be able to derive the identity $[\lambda F\, O!F \,\&\, x[\lambda z\, Fz]] = [\lambda F\, O!F \,\&\, xF]]$. Though the term on the right side of this identity isn't a core λ-expression, it nevertheless provably has a denotation; see (965.6).

Clearly, the metadefinition of *terms of the same type* (12) has to be modified. We'll say:

- τ and σ are terms of the same type iff for some type t, τ and σ are both terms of type t. (12)

We'll refer to this definition as (930) [12].

Again, we use the identity symbol = in both object language and metalanguage, with the expectation that, for any given occurrence, the context will make it clear whether object-theoretic identity or primitive metalinguistic identity is intended.

Finally, the following definitions can all be suitably adapted to the type-theoretic context:

- φ^{τ}_{α} and σ^{τ}_{α} (14)
- *substitutable at an occurrence* and *substitutable for* (15)
- *alphabetic variant* (16)

We'll use the same convention for referring to the type-theoretic versions of these definitions.

(**931**) **Remark:** Desiderata for the Definitions and Axioms of Typed Object Theory. Before we begin to develop definitions, state axioms, formulate a deductive system, and derive theorems for the language just defined, it is important to lay out the reasons why we may *not* axiomatize typed object theory simply by *typing all* of the axioms of the second-order theory. The language of typed object theory has far greater expressive power than that of second-order object theory. While some axioms of second-order object theory generalize to every appropriate type, others have to be restricted when typed, so as not to apply to all the new expressions in typed object theory. That is, certain axioms have to be adapted and limited so that they don't apply to expressions they weren't designed to govern.

To see specific examples of this, we begin with four key observations that must be kept in mind when formulating the type-theoretic versions of the axioms of second-order object theory:

(.1) Definite descriptions of type $\langle\,\rangle$ are both terms and formulas, and although all such descriptions have truth conditions (given that they are formulas), some fail to denote (i.e., fail to be significant), e.g., $\imath p(p \,\&\, \neg p)$ ('the proposition p such that both p and $\neg p$ are true').

(.2) A typed comprehension schema (which in second-order object theory asserted only the existence of abstract individuals) will now also assert, for each relational type t (i.e., for each type t where $t \neq i$), the existence of abstract relations of type t. An abstract relation of type t may both exemplify and encode properties of type $\langle t\rangle$. And an abstract relation of type t, together with $n-1$ objects ($n \geq 2$) having types $t_1,\ldots,t_{n-1}$, respectively, may exemplify and encode relations having one of the following types: $\langle t, t_1,\ldots,t_{n-1}\rangle$, $\langle t_1, t,\ldots,t_{n-1}\rangle$, ..., $\langle t_1,\ldots,t,t_{n-1}\rangle$, and $\langle t_1,\ldots,t_{n-1},t\rangle$. But the question arises, are any abstract relations themselves exemplified? Since there is no data suggesting that there are, and since the theoretical purpose of abstract entities of any type t is to encode and exemplify properties and relations of higher types, we shall stipulate that abstract relations are *unexemplified* and, when t is the type $\langle\,\rangle$ for propositions, that abstract propositions are *not true*.

(.3) Significant λ-expressions denote *ordinary* relations. From the point of view of typed object theory, the properties, relations, and propositions of second-order object theory are all *ordinary* relations. Thus, the purpose of λ-notation is to formulate complex expressions that would signify ordinary relations were they to have a significance. So, to preserve this intuition, it will be axiomatic that if a λ-expression is significant, it denotes an ordinary relation. We note here that when we analyze mathematics, we shall need to *introduce* a special group of *indexed* λ-expressions that are *defined* in terms of definite descriptions (of relational type) that signify abstract relations. While these will function as λ-expressions within the context of (truth in) a mathematical theory, outside those contexts they will be governed by the principles that govern their definientia and so function as definite descriptions (of relational type).

(.4) Formulas that aren't primitive constants, variables, or descriptions (i.e., formulas not in $Base^{\langle\rangle}$) denote *ordinary* propositions.

Each of these observations has consequences that we must bear in mind when formulating the axioms of typed object theory. We offer an extended discussion of each point in turn.

(.1) Let p be a variable ranging over the domain of propositions, i.e., ranging over the objects of type $\langle\,\rangle$. Note that the description $\imath p(p \& \neg p)$ is not just a term of type $\langle\,\rangle$, but also a formula (928). When this description appears as a term, as in the claims $O!\imath p(p\,\&\,\neg p)$ and $q = \imath p(p\,\&\,\neg p)$ (both of which are false, as we shall see), we read $\imath p(p\,\&\,\neg p)$ as: the proposition that is both true and false. Note that this description does not make an assertion when used as a term. Moreover, it provably denotes nothing. But when this description is used as a formula, such as when it stands alone or in the formula $\imath p(p\,\&\,\neg p) \equiv \exists p(p\,\&\,\neg p)$, then we read the formula as asserting: the proposition that is both true and false is true. So when in formula position, the expression $\imath p(p\,\&\,\neg p)$ is an assertion with truth conditions, despite the fact that it does not denote a proposition.[392] These facts lead to the (not necessarily independent) observations (.a) – (.f).

(.a) Consider the claim $\imath p(p\,\&\,\neg p)\downarrow$, in which $\imath p(p\,\&\,\neg p)$ figures as a term. This claim asserts that the proposition that is both true and false exists. No matter how *existence* is defined in type theory (see below), the claim $\imath p(p\,\&\,\neg p)\downarrow$ should imply $\exists p(p\,\&\,\neg p)$. Since the latter is provably false (any witness would be a contradiction), it should be provable that $\neg(\imath p(p\,\&\,\neg p)\downarrow)$. Given these facts, it follows immediately that: typed-versions of theorems (104.1) and (104.2) of second-order object theory cannot be theorems of typed-object theory. Theorem (104.1) asserts $\Pi^0\downarrow$, for every 0-ary relation term. Theorem (104.2) asserts $\varphi\downarrow$, for every formula φ. The corresponding claims of typed object theory are, respectively:

$\Pi^{\langle\rangle}\downarrow$, where $\Pi^{\langle\rangle}$ is any term of type $\langle\rangle$

$\varphi\downarrow$, for every formula φ

Neither of these can be theorems in typed object theory, for then it would be a theorem that $\imath p(p\,\&\,\neg p)\downarrow$. Thus, we have to weaken at least one of

[392] It might help to see that $\imath p(p\,\&\,\neg p)$ has well-defined truth conditions but doesn't denote a proposition if we think semantically for the moment. In the BNF (928), the line that stipulates which expressions are formulas tells us that descriptions are among the formulas in the *Base* expressions of type $\langle\rangle$ (which include the constants, variables, and descriptions of type $\langle\rangle$). Consider how our simultaneous definitions of denotation and truth in (929) imply that $\imath p(p\,\&\,\neg p)$ has well-defined truth conditions but doesn't denote a proposition. In clause T1 of (929), the definition of $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ states:

T1. Where $\varphi \in Base^{\langle\rangle}$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if $\exists \boldsymbol{p}(\boldsymbol{d}_{\mathcal{I},f}(\varphi) = \boldsymbol{p}\ \&\ \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{p}) = \boldsymbol{T})$.

So a definite description $\imath p(p\,\&\,\neg p)$ is true$_{\mathcal{I},f}$ at $\boldsymbol{w}$ if and only if it denotes$_{\mathcal{I},f}$ some proposition, say $\boldsymbol{p}$, and the exemplification extension$_{\boldsymbol{w}}$ of $\boldsymbol{p}$ is The True. But clause D3 governs the conditions under which $\imath p(p\,\&\,\neg p)$ has a denotation$_{\mathcal{I},f}$, namely, whenever there is a proposition uniquely satisfies the condition $p\,\&\,\neg p$. Since there can be no such proposition, the denotations conditions of $\imath p(p\,\&\,\neg p)$ imply that it doesn't denote a proposition. Nevertheless, $\imath p(p\,\&\,\neg p)$ still has well-defined truth conditions; these conditions just never obtain at any possible world, under any $\mathcal{I}$ and f, i.e., for any interpretation $\mathcal{I}$, assignment f, and world $\boldsymbol{w}$, it is not the case that $\imath p(p\,\&\,\neg p)$ is true$_{\mathcal{I},f}$ at $\boldsymbol{w}$. It is a logical falsehood.

The non-base formulas that are listed in the BNF (i.e., exemplification formulas, encoding formulas, $[\lambda\,\varphi]$, $\varphi \to \psi$, $\forall\alpha\varphi$, $\Box\varphi$, and $\mathcal{A}\varphi$) each have a separate clause in the definition of truth$_{\mathcal{I},f}$ at $\boldsymbol{w}$. Each clause therefore yields well-defined truth conditions even if φ is the non-denoting logical falsehood $\imath p(p\,\&\,\neg p)$.

the axioms or definitions from which $\Pi^0\!\downarrow$ and $\varphi\!\downarrow$ are derived.

In this case, the solution is to weaken a definition. Reconsider the proofs of (104.1) and (104.2). They both relied on the definition of proposition existence (20.3):

$$p\!\downarrow \equiv_{df} [\lambda x\, p]\!\downarrow \tag{20.3}$$

Clearly, this definition of proposition existence rests on the definition of property existence. But this dependency needn't be preserved in typed object theory, nor should it be. For the variable p functions as a metavariable in the above definition, and so the above definition would have the following instance:

$$\imath p(p \,\&\, \neg p)\!\downarrow \equiv_{df} [\lambda x\, \imath p(p \,\&\, \neg p)]\!\downarrow$$

Note that in the above instance, the definiens is true. By the type-theoretic version of axiom (39.2), the definiens is a core λ-expression: no variable bound by the λ occurs in encoding position in the matrix. So $[\lambda x\, \imath p(p \,\&\, \neg p)]$ signifies a property, and in particular, a (necessarily) empty property — no x can satisfy the matrix $\imath p(p \,\&\, \neg p)$. But if $[\lambda x\, \imath p(p \,\&\, \neg p)]$ signifies a property, then the above instance of the definition of proposition identity would incorrectly imply that $\imath p(p \,\&\, \neg p)\!\downarrow$.

So although definition (20.3) can't be preserved, a definition suitable to our type-theoretic framework is at hand. Where F has type $\langle\langle\,\rangle\rangle$, we may define:

$$p\!\downarrow \equiv_{df} \exists F\, Fp$$

This yields the instance:

$$\imath p(p \,\&\, \neg p)\!\downarrow =_{df} \exists F(F\imath p(p \,\&\, \neg p))$$

This definition lets us derive $\neg(\imath p(p \,\&\, \neg p)\!\downarrow)$.[393] And similarly, from the assumption that $\neg\imath p(p \,\&\, \neg p)\!\downarrow$, it would follow that $\neg F\imath p(p \,\&\, \neg p)$, by an instance of typed axiom (39.5.a) (see below). So $\forall F\neg F\imath p(p \,\&\, \neg p)$, i.e., $\neg\exists F(F\imath p(p \,\&\, \neg p))$.

In general, the revised definition ensures that an arbitrary term Π of type $\langle\,\rangle$ (and, thus, an arbitrary formula φ) is significant if and only if some property of propositions can be truly predicated of it by an exemplification predication.

[393] By Russell's theory of descriptions, $F\imath p(p \,\&\, \neg p)$ is equivalent to:

$$\exists p((p \,\&\, \neg p) \,\&\, \forall q((q \,\&\, \neg q) \to q = p) \,\&\, Fp)$$

But the negation of this claim is provable, since any witness would yield a contradiction. So $\neg F\imath p(p \,\&\, \neg p)$, where F is arbitrary. Hence $\forall F\neg F\imath p(p \,\&\, \neg p)$, i.e., $\neg\exists F(F\imath p(p \,\&\, \neg p))$. So by the definition of proposition identity just given in the text, $\neg(\imath p(p \,\&\, \neg p)\!\downarrow)$.

A similar result will be obtained for analogous terms, such as $\imath p(p \,\&\, \neg p \,\&\, Fx)$ – we should be able to prove that $\forall F\forall x\neg(\imath p(p \,\&\, \neg p \,\&\, Fx)\!\downarrow)$.

(.b) Moreover, now that we've seen both that the *formula* $\imath p(p\&\neg p)$ is false and that $\neg(\imath p(p \& \neg p)\downarrow)$ is derivable, it should be a theorem that $\neg\imath p(p \& \neg p)$. In general, any formula that fails to denote implies its own falsehood ('nonexistence implies falsehood'). These consequences are derivable from the typed version of (39.5.a) which, in the 0-ary case, will assert: $\tau^{\langle\rangle} \to \tau^{\langle\rangle}\downarrow$, where $\tau^{\langle\rangle}$ is any term of type $\langle\rangle$. Since formulas are terms of type $\langle\rangle$ (928), the claim $\varphi \to \varphi\downarrow$ is an instance (asserting 'truth implies existence') and so $\neg\varphi\downarrow \to \neg\varphi$ ('nonexistence implies falsehood') is just the contrapositive. Moreover, by RN, it follows that $\Box(\varphi \to \varphi\downarrow)$ and $\Box(\neg\varphi\downarrow \to \neg\varphi)$. Thus, in the case where φ is $\imath p(p \& \neg p)$, since it is provable that $\Box\neg\varphi\downarrow$, it follows that $\Box\neg\varphi$. Hence, it should be provable that $\imath p(p \& \neg p)$ is a necessary falsehood and equivalent to every other necessary falsehood.[394]

(.c) Theorem (111.1) of second-order object theory, which asserts $[\lambda\, \varphi] = \varphi$, where φ is any formula, cannot be a theorem of typed-object theory, since when φ is $\imath p(p \& \neg p)$, the instance $[\lambda\, \imath p(p \& \neg p)] = \imath p(p \& \neg p)$ would be a theorem. Typed object theory preserves the general principle that true identity statements imply that the terms flanking the identity sign have a denotation, i.e., typed versions of (107.1) and (107.2) should be theorems. So it would follow from $[\lambda\, \imath p(p \& \neg p)] = \imath p(p \& \neg p)$ and (107.2) that $\imath p(p \& \neg p)\downarrow$. Moreover, if τ is a significant proposition term of type $\langle\rangle$ but denotes an abstract proposition, then $[\lambda\, \tau] = \tau$ will also fail – by hypothesis, the right side denotes something abstract, but by (.3), $[\lambda\, \tau]$ will denote an *ordinary* proposition, since it will be significant.

Thus, to prove a weakened version of (111.1), we have to weaken one of the axioms on which the theorem $[\lambda\, \varphi] = \varphi$ depends. Indeed, the axiom that needs to be weakened is η-Conversion, which holds only for ordinary relations. See (935.27) below. Consequently, the relevant theorem of typed object theory will be: $[\lambda\, \varphi] = \varphi$, provided φ is not a basic formula, i.e., not a formula in $Base^{\langle\rangle}$. This is proved in (940.3a) below.

(.d) While the term $\imath p(p \& \neg p)$ fails to be significant, the term $[\lambda\, \imath p(p \& \neg p)]$ ("that the proposition that is both true and false is true") is nevertheless significant – we'll see later that it denotes an ordinary, but false, proposition. That is, it should be a theorem that:

(.i) $[\lambda\, \imath p(p \& \neg p)]\downarrow$

Indeed, this is axiomatic, since it will be an instance of the type-theoretic version of (39.2) – no variable bound by the λ occurs in encoding position in the matrix.

[394] For example, it should be derivable that $\Box(\imath p(p \& \neg p) \equiv \forall p(p \& \neg p))$.

Intuitively, the λ in a λ-expression is a constructor for logically possible exemplification patterns, i.e., exemplification patterns that don't lead to contradiction, even if defined in terms of encoding formulas. So a λ-expression in which the λ binds no variables is a constructor for logically possible truth/falsehood patterns. The expression $[\lambda\,\varphi]$ can be read intuitively as *that φ is true* and thus signifies a proposition, namely, the 0-ary exemplification of φ. As such, $[\lambda\,\varphi]\!\downarrow$ is true even when $\neg(\varphi\!\downarrow)$. Thus, by adding λ to $\imath p(p \,\&\, \neg p)$ to form the term $[\lambda\,\imath p(p \,\&\, \neg p)]$, we're constructing an expression that asserts a (necessarily) false proposition, namely, that the true and false proposition is true.

The general conclusions to draw from this are:

(.ii) since every formula φ is true or false, the claim $[\lambda\,\varphi]\!\downarrow$ is axiomatic, and

(.iii) $[\lambda\,\varphi]\!\downarrow \equiv \varphi\!\downarrow$ is not a theorem, since the empty description $\imath p(p \,\&\, \neg p)$ causes this biconditional to fail.[395]

By contrast, the 0-ary case of β-Conversion, namely $[\lambda\varphi] \equiv \varphi$ can be preserved. For example, $[\lambda\,\imath p(p \,\&\, \neg p)] \equiv \imath p(p \,\&\, \neg p)$ is true because both sides are false.

(.e) The following should hold:

(.i) $[\lambda\,\varphi]\!\downarrow \rightarrow \varphi\!\downarrow$, provided that φ is not a description

either because the following is axiomatic:

(.ii) $\varphi\!\downarrow$, provided that φ is not a description

in which case (.i) would follow from (.ii), or because (.i) itself is axiomatic. In (.3.g) below, we consider an axiom that implies (.ii).

Note that (.e.ii) *allows* for cases of descriptions of type $\langle\,\rangle$ of the form $\imath p\varphi$ that are provably significant under the new definition of proposition existence discussed in (.a) above. For example, we should be able to prove $q = \imath p(p = q)$ and thus that $\imath p(p = q)\!\downarrow$. Moreover, where $A!$ is the property of *being abstract* having type $\langle\langle\,\rangle\rangle$ and F is a variable of type $\langle\langle\,\rangle\rangle$, we should be able to prove $\imath p(A!p \,\&\, \forall F(pF \equiv \varphi))\!\downarrow$, for any φ in which p doesn't occur free. See below for further discussion of the conditions under which we should be able to prove $\varphi\!\downarrow$.

(.f) Theorem (111.4) of second-order object theory, i.e., $\varphi = \varphi'$ for any alphabetic variants φ and φ', cannot be a theorem of typed object theory, for when φ is $\imath p(p \,\&\, \neg p)$ and φ' is $\imath q(q \,\&\, \neg q)$, it would be a theorem that $\imath p(p \,\&\, \neg p) = \imath q(q \,\&\, \neg q)$. We cannot accept such a consequence

[395] In second-order object theory, $[\lambda\,\varphi]\!\downarrow \equiv \varphi\!\downarrow$ follows immediately from the fact that $[\lambda\,\varphi]\!\downarrow$ is an axiom (39.2), and $\varphi\!\downarrow$ is a theorem (104.2). But in typed object theory, $[\lambda\,\varphi]\!\downarrow \equiv \varphi\!\downarrow$ holds only when $\varphi\!\downarrow$. So when φ is a canonical description of an abstract proposition, of the form $\imath p(A!p \,\&\, \forall F(pF \equiv \varphi))$, then the biconditional holds, since both sides are true.

while retaining the general principle that true identity statements imply that the terms flanking the identity sign have a denotation, i.e., (107.1) and (107.2). Thus, we have to weaken at least one of the axioms upon which the theorem $\varphi = \varphi'$ depends. But we should be able to derive $\varphi\!\downarrow \to (\varphi = \varphi')$ and, when φ is not a description, that $\varphi = \varphi'$.

(.2) Intuitively, any relation whose existence is asserted by an instance of the typed comprehension principle for *abstract* relations is to be conceived solely as a reified encoding pattern. For such relations, we stipulate that they are unexemplified. Here is a more exact description of this stipulation.

(.a) Abstract binary relations among individuals are unexemplified by any individuals, ordinary or abstract. For example, where x is a variable of type $\langle i,i\rangle$ that doesn't occur free in φ, $A!$ is a defined term of type $\langle\langle i,i\rangle\rangle$, and F is a variable of type $\langle\langle i,i\rangle\rangle$, the abstract binary relations among individuals asserted by instances of the comprehension axiom $\exists x(A!x \,\&\, \forall F(xF \equiv \varphi))$ are not exemplified. Abstract, binary relations reify patterns of properties of binary relations among individuals and there is no pre-theoretic reason to suppose that such abstract relations are exemplified. Thus, when we identify the membership relation of ZF ($\in_{\mathrm{ZF}}$) as an abstract relation, it will only encode the properties of relations *exemplified by* $\in$ *in* ZF and it will exemplify properties and stand in relations as well. But it will not be exemplified by any individuals.

In general, where $\langle t_1,\ldots,t_n\rangle$ is any relational type, F is a variable of that type, $A!$ is a term of type $\langle\langle t_1,\ldots,t_n\rangle\rangle$ (so that $A!F$ asserts that F is abstract), and $x_1,\ldots,x_n$ are variables having types $t_1,\ldots,t_n$, respectively, then we shall want, as an axiom, that:

$$A!F \to \neg\exists x_1 \ldots \exists x_n F x_1 \ldots x_n \qquad (n \geq 0)$$

(.b) Similarly, since the 0-ary case of predication is truth, abstract propositions are always false. Abstract propositions reify patterns of properties of propositions and there is no pre-theoretic reason to suppose that such abstract propositions are true. So where p is a variable of type $\langle\,\rangle$, $A!$ is a term of type $\langle\langle\,\rangle\rangle$ (so that $A!p$ asserts that p is abstract), we have, as the 0-ary case of the axiom in (.2.a), that: $A!p \to \neg p$.

(.3) To ensure that λ-expressions denote ordinary relations when they are significant, consider the following axiom schema:

(.a) Let $x_1,\ldots,x_n$ $(n \geq 0)$ be variables having type $t_1,\ldots,t_n$ respectively, $O!$ have type $\langle\langle t_1,\ldots,t_n\rangle\rangle$, and $[\lambda x_1 \ldots x_n \varphi]$ have type $\langle t_1,\ldots,t_n\rangle$, so that the formula $O![\lambda x_1 \ldots x_n\, \varphi]$ asserts that $[\lambda x_1 \ldots x_n\, \varphi]$ is an ordinary relation. Then the closures of the following are axioms:

$$[\lambda x_1 \dots x_n\, \varphi]\downarrow \to O![\lambda x_1 \dots x_n\, \varphi] \qquad (n \geq 0)$$

This holds only for λ-expressions and does not apply to formulas; $\varphi\downarrow \to O!\varphi$ is not an instance of this schema. Canonical descriptions of abstract propositions of the form $\imath p(A!p \,\&\, \forall F(pF \equiv \varphi))$ are provably significant but provably do not denote ordinary propositions. Such descriptions would be examples of a formula φ such that $\varphi\downarrow$ and $\neg O!\varphi$. (More on this below.)

Now β-Conversion can be formulated in the usual way, as restricted to significant λ-expressions. However, in contrast to second-order object theory, we specify that $n \geq 0$ in the statement of β-Conversion (in second-order object theory, the 0-ary case of β-Conversion is derived), so that β-Conversion becomes:

(.b) Where $x_1, \dots, x_n$ ($n \geq 0$) are distinct variables having types $t_1, \dots, t_n$, respectively, then the closures of the following are axioms:

$$[\lambda x_1 \dots x_n\, \varphi]\downarrow \to ([\lambda x_1 \dots x_n\, \varphi]x_1 \dots x_n \equiv \varphi) \qquad (n \geq 0)$$

Note that this holds even in the 0-ary cases where φ is $\imath p(p \,\&\, \neg p)$ and where φ is a canonical description of an abstract proposition. For when $n = 0$ and φ is $\imath p(p \,\&\, \neg p)$, we have:

$$[\lambda\, \imath p(p \,\&\, \neg p)]\downarrow \to ([\lambda\, \imath p(p \,\&\, \neg p)] \equiv \imath p(p \,\&\, \neg p))$$

Since the antecedent holds, the consequent is true. But the consequent is true because both sides are (necessarily) false. Similarly, consider $\imath p(A!p \,\&\, \forall F(pF \equiv \varphi))$ (i.e., any *canonical* description of an *abstract* proposition). Then:

$$\begin{aligned}&[\lambda\, \imath p(A!p \,\&\, \forall F(pF \equiv \varphi))]\downarrow \to \\ &\quad ([\lambda\, \imath p(A!p \,\&\, \forall F(pF \equiv \varphi))] \equiv \imath p(A!p \,\&\, \forall F(pF \equiv \varphi)))\end{aligned}$$

Again, the antecedent is true, and the consequent is true because both sides of the biconditonal are (necessarily) false. The difference between this case and the last is that the description $\imath p(p \,\&\, \neg p)$ is (necessarily) false because it fails to denote a proposition, but the description $\imath p(A!p \,\&\, \forall F(pF \equiv \varphi))$ is (necessarily) false because it denotes an abstract proposition – abstract relations aren't exemplified and, in the 0-ary case, abstract propositions aren't true.

Note that the converse of (.a), i.e., $O![\lambda x_1 \dots x_n\, \varphi] \to [\lambda x_1 \dots x_n\, \varphi]\downarrow$, is also axiomatic; by the typed version of (39.5.a), true exemplification formulas imply that the primary terms in the formula are significant. So we can derive the following biconditional:

(.c) $[\lambda x_1 \dots x_n\, \varphi]\downarrow \equiv O![\lambda x_1 \dots x_n\, \varphi]$ $\qquad (n \geq 0)$

Hence, by the type-theoretic version of β-Conversion formulated in (.b), we can say that every *ordinary* relation, of whatever type, can be β-converted.

Let's consider an interesting case of (.a), where φ is the propositional variable p, x is of any type t, $[\lambda x\, p]$ therefore has type $\langle t\rangle$, and $O!$ has type $\langle\langle t\rangle\rangle$. The following is a universal closure of a unary case of (.a) and so is axiomatic:

(.d) $\forall p([\lambda x\, p]{\downarrow} \to O![\lambda x\, p])$

Consider again $\imath p(A!p \,\&\, \forall F(pF \equiv \varphi))$, but where x doesn't occur free in φ. Since descriptions of this form have type $\langle\,\rangle$ and are guaranteed to be significant by comprehension and identity principles, we may instantiate them into (.d) to obtain:

$$[\lambda x\, \imath p(A!p \,\&\, \forall F(pF \equiv \varphi))]{\downarrow} \to O![\lambda x\, \imath p(A!p \,\&\, \forall F(pF \equiv \varphi))].$$

Note that the antecedent is derivable, for $[\lambda x\, p]{\downarrow}$ is axiomatic and since this holds for all p, it holds for $\imath p(A!p \,\&\, \forall F(pF \equiv \varphi))$ given that x doesn't occur free in φ. Since the antecedent is derivable, it follows that:

$$O![\lambda x\, \imath p(A!p \,\&\, \forall F(pF \equiv \varphi))]$$

Thus, we have two reasons to suppose $[\lambda x\, \imath p(A!p \,\&\, \forall F(pF \equiv \varphi))]{\downarrow}$ — one reason is that the fact displayed immediately above implies it (by way of (.c) or by way of the fact that a true exemplification formula implies that its primary terms are significant), and the second reason is that no variable bound by the λ occurs in encoding position in the matrix (recall that we're considering the case where x doesn't occur free in φ).

Hence, by β-Conversion:

$$[\lambda x\, \imath p(A!p \,\&\, \forall F(pF \equiv \varphi))]x \equiv \imath p(A!p \,\&\, \forall F(pF \equiv \varphi))$$

Since $\imath p(A!p \,\&\, \forall F(pF \equiv \varphi))$ is an abstract proposition, it is always false, by (.2.b) above. So the right side of the above equivalence is false and, hence, so is the left side, for any x. Thus, although $[\lambda x\, \imath p(A!p \,\&\, \forall F(pF \equiv \varphi))]$ provably exists, it is provably unexemplified!

Note further that in the 0-ary case of (.a), we have:

(.e) $[\lambda\, \varphi]{\downarrow} \to O![\lambda\, \varphi]$

Since the type-theoretic version of axiom (39.2) will stipulate that $[\lambda\, \varphi]{\downarrow}$ (this is a core λ expression), it follows that $O![\lambda\, \varphi]$.

Next, we turn to η-Conversion, which holds for elementary λ-expressions in which the 'head' relation is ordinary:

(.f) Where F is a variable of $\langle t_1,\ldots,t_n\rangle$, $O!$ has type $\langle\langle t_1,\ldots,t_n\rangle\rangle$ (so that $O!F$ asserts that F is ordinary), and $x_1,\ldots,x_n$ are variables of type $t_1,\ldots,t_n$, respectively, η-Conversion becomes:

$$O!F \to ([\lambda x_1 \ldots x_n\, Fx_1 \ldots x_n] = F)$$

In other words, if F is an ordinary relation, then *being* $x_1,\ldots,x_n$ *such that* $Fx_1\ldots x_n$ is identical to F. In the 0-ary case, this becomes:

$$O!p \rightarrow ([\lambda\, p] = p)$$

In other words, whenever p is an ordinary proposition, *that*-p is identical to p. Since $\neg\imath p(p \,\&\, \neg p)\downarrow$, we can't instantiate $\imath p(p \,\&\, \neg p)$ for p in the above, and we can't derive the known falsehood $[\lambda\, \imath p(p \,\&\, \neg p)] = \imath p(p \,\&\, \neg p)$. However, since $[\lambda\, \imath p(p \,\&\, \neg p)]\downarrow$, we can instantiate it in the above to obtain:

$$O![\lambda\, \imath p(p \,\&\, \neg p)] \rightarrow ([\lambda\, [\lambda\, \imath p(p \,\&\, \neg p)]] = [\lambda\, \imath p(p \,\&\, \neg p))]$$

Note that none of the above implies either of:

$$\varphi\downarrow \rightarrow O!\varphi$$

$$\forall p(p\downarrow \rightarrow O!p)$$

These are subject to counterexample; as we noted above, let φ be a canonical description of an *abstract* proposition whose existence is guaranteed by comprehension and identity principles. To take a particular example, consider the abstract proposition that encodes exactly one property, namely, the property of being a proposition q that is true iff not true:

$$\imath p(A!p \,\&\, \forall F(pF \equiv F = [\lambda q\, (q \equiv \neg q)]))$$

This is a formula and a term that denotes a proposition. So it can be used to form an instance of $\varphi\downarrow \rightarrow O!\varphi$ and can be instantiated into $\forall p(p\downarrow \rightarrow O!p)$. Both the instance and the result of the instantiation would be false, since the description is significant but, by the theory of descriptions, denotes an abstract (i.e., not ordinary) proposition.

Note also that none of the above implies the result:

$$[\lambda\, \imath p(A!p \,\&\, \forall F(pF \equiv F = [\lambda q\, (q \equiv \neg q)]))] = \imath p(A!p \,\&\, \forall F(pF \equiv F = [\lambda q\, (q \equiv \neg q)]))$$

This identity will be provably false. The left side of the identity denotes an ordinary proposition, while the right side denotes an abstract proposition. To see why the left side denotes an ordinary proposition, recall that (i) λ-expressions in which the λ doesn't bind a variable are significant and (ii) significant λ-expressions, by (.3.a) and (.3.e) above, denote ordinary propositions. Clearly, the example displayed immediately above shows that, for arbitrary formulas φ, the following should not be theorems:

$$[\lambda\, \varphi]\downarrow \rightarrow ([\lambda\, \varphi] = \varphi)$$

$$\varphi\downarrow \rightarrow ([\lambda\, \varphi] = \varphi)$$

(.4) In general, we shall need, as an axiom:

$O!\varphi$, provided φ is not a constant of type $\langle\rangle$, a variable of type $\langle\rangle$, or a description of type $\langle\rangle$

We exclude the expressions in $Base^{\langle\rangle}$ for the following reasons. Constants and variables of type $\langle\rangle$ are excluded because some denote abstract propositions or take abstract propositions as a value. Empty descriptions are excluded because they don't denote anything and so don't denote something that is ordinary. And there are canonical descriptions of the form $\imath p(A!p \,\&\, \forall F(pF \equiv \varphi))$ that denote abstract propositions and so denote something that isn't ordinary.

By constrast, formulas not in $Base^{\langle\rangle}$ denote ordinary propositions. So the axiom described as (.4) above is related to the desideratum described in (.1.e.ii), namely, that it should be a theorem that $\varphi\downarrow$, provided φ is not a description. (.4) implies (.1.e.ii), as follows: if φ is a constant or variable of type $\langle\rangle$, then $\varphi\downarrow$ by the type-theoretic version of (39.2); and if φ is neither a constant, a variable nor a description, then by (.4), $O!\varphi$, and so by the type-theoretic version of (39.5.a), $\varphi\downarrow$.[396]

(**932**) **Remark:** On Definitions in Typed Object Theory. The theory of definitions developed in second-order object theory works well for typed object theory. But a few observations are in order.

Often we don't need to worry about non-denoting formulas in a definitions-by-$\equiv$. For example, consider:

$$\varphi \,\&\, \psi \equiv_{df} \neg(\varphi \to \neg\psi)$$

If we let φ be the non-denoting formula $\imath p(p \,\&\, \neg p)$, we would have:

$$\imath p(p \,\&\, \neg p) \,\&\, \psi \equiv_{df} \neg(\imath p(p \,\&\, \neg p) \to \neg\psi)$$

This doesn't create any special problems. It will be provable that:

$$\neg(\imath p(p \,\&\, \neg p) \,\&\, \psi)$$

on the grounds that $\neg\imath p(p \,\&\, \neg p)$. Even though $\imath p(p \,\&\, \neg p)$ doesn't denote, it is a formula that has truth conditions that are always false.

Indeed, we can use a description like $\imath p(p\&\neg p)$ in both definitions-by-$\equiv$ and definitions-by-$=$, such as in the following:

(ϑ) $q_1 \equiv_{df} \imath p(p \,\&\, \neg p)$

(ξ) $q_1 =_{df} \imath p(p \,\&\, \neg p)$

[396] If we were to try to prove (.4) instead of taking it as an axiom, we would have to traverse the BNF for φ by skipping the Base formulas (constants, variables, and descriptions). But it is not clear how we could even establish the base cases, i.e., show that exemplification and encoding formulas denote ordinary propositions, no matter whether true or false. So (.4) will be an axiom.

By our theory of definitions-by-equivalence, the inferential role of (ϑ) is to introduce $\Box(q_1 \equiv \imath p(p \& \neg p))$ as an axiom. So the new expression q_1 is introduced as necessarily equivalent to a necessary falsehood and so is necessarily false. We can prove that $\neg q_1$ from the theorem that $\neg \imath p(p \,\&\, \neg p)$. By constrast, in the theory of definitions-by-identity, the inferential role of (ξ) is to introduce the axiom:

$$(\imath p(p \,\&\, \neg p)\!\downarrow \to (q_1 = \imath p(p \,\&\, \neg p))) \,\&\, (\neg(\imath p(p \,\&\, \neg p)\!\downarrow) \;\to \neg(q_1\!\downarrow))$$

Since $\neg \imath p(p \,\&\, \neg p)\!\downarrow$ will be provable, (ξ) therefore implies $\neg(q_1\!\downarrow)$. One must therefore deploy (ϑ) instead of (ξ) if it is important to be able to substitute q_1 for $\imath p(p \,\&\, \neg p)$ is some context in which one or the other is a subformula. Of course, to simplify a proof, one might introduce q_1 as a *mere notational abbreviation* of the formula $\imath p(p \,\&\, \neg p)$ and substitute one for the other in a proof.

However, in other cases, we have to take precautions against formulas that fail to denote. For example consider what happens if we instantiate $\imath p(p \,\&\, \neg p)$ into a definition such as:

$$Rigid(\varphi) \equiv_{df} \varphi \to \Box\varphi$$

This would yield the instance:

$$Rigid(\imath p(p \,\&\, \neg p)) \equiv_{df} \imath p(p \,\&\, \neg p) \to \Box \imath p(p \,\&\, \neg p)$$

Since we know $\neg \imath p(p \,\&\, \neg p)$ (given that non-existence implies falsehood) then the antecedent of the definiens fails, making the definiens true. Thus we would have a proof of $Rigid(\imath p(p \,\&\, \neg p))$, contrary to the garbage-in, garbage-out principle. So we either have to reformulate the definition as:

$$Rigid(\varphi) \equiv_{df} \varphi\!\downarrow \,\&\, (\varphi \to \Box\varphi)$$

or as:

$$Rigid(p) \equiv_{df} p\!\downarrow \,\&\, (p \to \Box p)$$

where we use the standard convention for definitions whereby object language variables function as metavariables.

Note, finally, that the observations in Remark (109) do not apply to typed object theory. In Remark (109), we observed there (a) that since $\varphi\!\downarrow$ is a theorem, every definiens is significant and (b) that since identities between formulas yield equivalences (108), any definition-by-equivalence could be recast as a definition-by-identity. But we also explained why we didn't exercise the option of eliminating definitions-by-equivalence in second-order object theory, since the move would eliminate the hyperintensionality built into the system. But, as noted in footnote 146, not every formula is significant in typed object theory. A definition by equivalence yields an identity only if the definiens is significant. So, in typed object theory, we don't have a viable option of eliminating definitions by equivalence in favor of definitions by identity, even if we had wanted to ignore the demands of hyperintensionality.

15.2 Definitions

(**933**) **Definitions:** Definitions for Typed Object Theory. We now introduce some defined notions into our typed language. First, we import all the definitions, from (18), of the classical connectives (&, $\vee$, and $\equiv$), the existential quantifier ($\exists$), and the possibility operator ($\Diamond$). So let (.1) – (.5) be the definitions of these five notions, exactly as stated in (18.1) – (18.5):

(.1) $\varphi \,\&\, \psi \;\equiv_{df}\; \neg(\varphi \rightarrow \neg\psi)$

(.2) $\varphi \vee \psi \;\equiv_{df}\; \neg\varphi \rightarrow \psi$

(.3) $\varphi \equiv \psi \;\equiv_{df}\; (\varphi \rightarrow \psi) \,\&\, (\psi \rightarrow \varphi)$

(.4) $\exists\alpha\varphi \;\equiv_{df}\; \neg\forall\alpha\neg\varphi$

(.5) $\Diamond\varphi \;\equiv_{df}\; \neg\Box\neg\varphi$

We next define existence. We first define existence for the lowest types i and $\langle\,\rangle$, and then define it for relational types $\langle t_1,\ldots,t_n\rangle$ ($n \geq 1$). Let x be a variable of type i, and F be a variable of type $\langle i\rangle$, and suppose x and F function as metavariables. Then we say that (.6.a) *x exists* just in case x exemplifies some property:

(.6.a) $x{\downarrow} \equiv_{df} \exists F\, Fx$

Now let p be a variable of type $\langle\,\rangle$ and F be a variable of type $\langle\langle\,\rangle\rangle$, and suppose p and F function as metavariables. Then we say (.6.b) *p exists* just in case p exemplifies some property:

(.6.b) $p{\downarrow} \equiv_{df} \exists F\, Fp$

Thus, the definitions for *existence* of the lowest type entities, i and $\langle\,\rangle$, are analogous.

Now for any types $t_1,\ldots,t_n$, let F be a variable of type $\langle t_1,\ldots,t_n\rangle$ ($n \geq 1$), and $x_1,\ldots,x_n$ be variables of types $t_1,\ldots,t_n$, respectively, and let the variables function as metavariables. We then say (.6.c) *F exists* just in case there are objects $x_1,\ldots,x_n$ that encode F:

(.6.c) $F{\downarrow} \equiv_{df} \exists x_1 \ldots \exists x_n (x_1 \ldots x_n F)$ $\qquad (n \geq 1)$

To introduce the definitions for identity, we first define, for any type t, the properties of *being ordinary* and *being abstract* having the type $\langle t\rangle$, so that both may be predicated of objects of type t. For this definition, let $E!$ be the distinguished constant of type $\langle t\rangle$ that denotes the primitive property of *being concrete* that can be predicated of objects of type t, and let x be a variable of type t. Then (.7) *being ordinary* is, by definition, *being an x such that x is possibly concrete*, and (.8) *being abstract* is, by definition, *being an x such that x couldn't possibly be concrete*:

(.7) $O! =_{df} [\lambda x \, \Diamond E!x]$

(.8) $A! =_{df} [\lambda x \, \neg\Diamond E!x]$

Next we define formulas of the form $\alpha^t = \beta^t$, where t is any type. The definition can be given in complete generality by providing definientia for the following cases:

- $x = y$, where x and y have type i
- $F = G$, where F and G have type $\langle t\rangle$, where t is any type
- $F = G$, where F and G have type $\langle t_1,\ldots,t_n\rangle$ $(n \geq 2)$ and $t_1,\ldots,t_n$ are any types
- $p = q$, where p and q have type $\langle\,\rangle$

For the first case, let x and y be variables of type i, $O!$ and $A!$ have type $\langle i\rangle$, and F be a variable of type $\langle i\rangle$. Then we define:

(.9) $x = y \equiv_{df} (O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)) \vee (A!x \,\&\, A!y \,\&\, \Box\forall F(xF \equiv yF))$

I.e., individuals x and y are *identical* if and only if either (a) they are both ordinary individuals and necessarily exemplify the same properties of type $\langle i\rangle$, or (b) they are both abstract individuals and necessarily encode the same properties of type $\langle i\rangle$.

For the next case, let t be any type, F and G be variables of type $\langle t\rangle$, $O!$ and $A!$ have type $\langle\langle t\rangle\rangle$, x be a variable of type t, and $\mathcal{H}$ be a variable of type $\langle\langle t\rangle\rangle$. Then we define identity for properties having type $\langle t\rangle$ as follows:

(.10) $F = G \equiv_{df} (O!F \,\&\, O!G \,\&\, \Box\forall x(xF \equiv xG)) \vee (A!F \,\&\, A!G \,\&\, \Box\forall\mathcal{H}(F\mathcal{H} \equiv G\mathcal{H}))$

I.e., F and G are *identical* if and only if either (a) F and G are both ordinary properties that are necessarily encoded by the same objects of type t, or (b) F and G are both abstract properties that necessarily encode the same properties having type $\langle\langle t\rangle\rangle$.

For the next case, let $t_1,\ldots,t_n$ be any types $(n \geq 2)$, F and G be variables of type $\langle t_1,\ldots,t_n\rangle$, $O!$ and $A!$ have type $\langle\langle t_1,\ldots,t_n\rangle\rangle$, $x_1,\ldots,x_n$ be variables of type $t_1,\ldots,t_n$, respectively, and $\mathcal{H}$ be a variable of type $\langle\langle t_1,\ldots,t_n\rangle\rangle$. Then we define identity for relations having type $\langle t_1,\ldots t_n\rangle$ as follows:

(.11) $F = G \equiv_{df}$

$$\begin{aligned}&O!F \,\&\, O!G \,\&\, \forall x_2 \ldots \forall x_n([\lambda x_1\, Fx_1 \ldots x_n] = [\lambda x_1\, Gx_1 \ldots x_n]) \,\&\\ &\forall x_1 \forall x_3 \ldots \forall x_n([\lambda x_2\, Fx_1 \ldots x_n] = [\lambda x_2\, Gx_1 \ldots x_n]) \,\&\ldots\&\\ &\forall x_1 \ldots \forall x_{n-1}([\lambda x_n\, Fx_1 \ldots x_n] = [\lambda x_n\, Gx_1 \ldots x_n]) \quad \vee\\ &A!F \,\&\, A!G \,\&\, \Box\forall\mathcal{H}(F\mathcal{H} \equiv G\mathcal{H})\end{aligned}$$

I.e., F and G are *identical* just in case either (a) F and G are both ordinary relations and such that every way of projecting F and G onto any $n-1$ objects of the right type yields identical properties, or (b) F and G are both abstract relations that necessarily encode the same properties with type $\langle\langle t_1, \ldots, t_n\rangle\rangle$.

For the final case, let p and q be variables of type $\langle\,\rangle$, $O!$ and $A!$ have type $\langle\langle\,\rangle\rangle$, x be a variable of type i, and $\mathcal{H}$ be a variable of type $\langle\langle\,\rangle\rangle$. Then we define identity for propositions as follows:

(.12) $p = q \equiv_{df} (O!p \,\&\, O!q \,\&\, [\lambda x\, p] = [\lambda x\, q]) \vee (A!p \,\&\, A!q \,\&\, \Box\forall\mathcal{H}(p\mathcal{H} \equiv q\mathcal{H}))$

I.e., p and q are *identical* just in case either (a) p and q are both ordinary propositions for which *being an individual such that* p is identical to *being an individual such that* q, or (b) p and q are both abstract propositions that necessarily encode the same properties with type $\langle\langle\,\rangle\rangle$.

(**934**) **Remark:** Observations on Identity in Typed Object Theory. The above definitions of identity differ somewhat from those of earlier publications on object theory. In Zalta 1982 (301–2), 2000b (228), and 2020 (71), we used the following definition, where x and y are variables of type t, $O!$ and $A!$ are constants of type $\langle t\rangle$, and F is a variable of type $\langle t\rangle$, for any type t:

(A) $x = y \equiv_{df} (O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)) \vee (A!x \,\&\, A!y \,\&\, \Box\forall F(xF \equiv yF))$

Notice that here we defined the identity of ordinary entities of any type t in terms of the $\langle t\rangle$-properties they exemplify. Though (A) works just fine (see below), it implies that ordinary objects of relational type $t \neq i$ are identical if they necessarily exemplify the same properties that have type $\langle t\rangle$. But (A) isn't a generalization of the second-order definition of relation identity, for the relations of second-order object theory are, from the point of view of typed object theory, *ordinary* relations. To see why the first disjunct of (A) isn't a generalization of the second-order definition, consider the case of properties of type $\langle i\rangle$. In second-order object theory, properties F and G are identical just in case, necessarily, they are encoded by the same individuals. But (A) stipulates that if F and G are ordinary properties of type $\langle i\rangle$, they are identical just in case they exemplify the same properties that have type $\langle\langle i\rangle\rangle$. Thus, (A) fails to preserve the insight that properties F and G are predicable entities; by contrast, their identity conditions in second-order object theory are defined in terms of their role as predicable entities.

Nevertheless, it will be provable that all and only the ordinary properties that are necessarily encoded by the same objects are ordinary properties that exemplify the same properties of properties. See (967.2) below for a proof of:

$$(O!F \,\&\, O!G \,\&\, \Box\forall x(xF \equiv yG)) \equiv (O!F \,\&\, O!G \,\&\, \Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G))$$

Thus, the definition of identity given in Zalta 1982 (301–2), 2000b (228), and 2020 (71) aren't incorrect – they at least imply the insights on which the current definition is based, since they imply that ordinary properties that necessarily exemplify the same properties are necessarily encoded by the same objects.

It should also be mentioned that in Zalta 1983 (121, 124) and 1988 (241–2), we defined identity for higher-order abstracta in terms that didn't sufficiently distinguish the ordinary and abstract objects of types $t \neq i$. That is, in these works, we assumed that if F and G are *any* higher-order properties of type $\langle t\rangle$, for some type t, then they are identical if and only if they are necessarily encoded by the same objects of type t, i.e., we stipulated that:

(B) $F = G \;\equiv_{df}\; \Box\forall x^t(xF \equiv xG)$

Relation identity was then defined in terms of (B). But these definitions of identity for properties of type $\langle t\rangle$ don't sufficiently distinguish the identity conditions for ordinary properties of type $\langle t\rangle$ (namely, when they are necessarily encoded by the same objects of type t) from the identity conditions for abstract properties of type $\langle t\rangle$ (namely, when they necessarily encode the same properties of type $\langle\langle t\rangle\rangle$). So the definitions from these two works should no longer be used.

In conclusion, the present formulation of the definition of identity in (933.9) is based on the idea that at each relational type, the *ordinary* properties and relations of that type have identity conditions defined in terms of the objects that encode them, whereas the *abstract* properties and relations of that type have identity conditions defined in terms of the higher-type properties that they encode. Thus, the identity conditions for ordinary properties, relations, and propositions are specified in terms that recognize their fundamental character as predicable entities, while the identity conditions for abstract properties, relations, and propositions are specified in terms that recognize their fundamental character as encoders.

15.3 Axioms

If we keep the issues raised in Remark (931) in mind, then most of the axioms can be typed in the usual way, with only a few to address the increased expressive power.

(**935**) **Axioms:** Typing the Axioms. In what follows, when we assert the closures of a schema as an axiom, we thereby assert the closures of all the instances of the schema as axioms.

Negations and Conditionals. The axioms for negation and conditionals are the same as those of second-order object theory. Where φ, ψ, and χ are any formulas, the closures of following are all necessary axioms:

(.1) $\varphi \to (\psi \to \varphi)$

(.2) $(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi))$

(.3) $(\neg\varphi \to \neg\psi) \to ((\neg\varphi \to \psi) \to \varphi)$

Quantification and Logical Existence. The axioms governing quantification and logical existence are typed versions of the axioms used in the second-order case. Let $t, t_1, \ldots, t_n$ be any types, and let $\tau, \tau_1, \ldots, \tau_n$ be any terms respectively having these types. Then where α is any variable of type t and Π is any term with type $\langle t_1, \ldots, t_n\rangle$ $(n \geq 0)$, we assert the closures of the following as axioms:

(.4) $\forall\alpha\varphi \to (\tau\downarrow \to \varphi^{\tau}_{\alpha})$, provided τ is substitutable for α in φ

(.5) $\tau\downarrow$, whenever τ is either a primitive constant, a variable, or a core λ-expression

(.6) $\forall\alpha(\varphi \to \psi) \to (\forall\alpha\varphi \to \forall\alpha\psi)$

(.7) $\varphi \to \forall\alpha\varphi$, provided α doesn't occur free in φ

(.8) (a) $\Pi\tau_1 \ldots \tau_n \to (\Pi\downarrow \,\&\, \tau_1\downarrow \,\&\, \ldots \,\&\, \tau_n\downarrow)$ $(n \geq 0)$
(b) $\tau_1 \ldots \tau_n\Pi \to (\Pi\downarrow \,\&\, \tau_1\downarrow \,\&\, \ldots \,\&\, \tau_n\downarrow)$ $(n \geq 1)$

Note that (.8.a) is the counterpart of (39.5.a) and reduces to $\Pi^{\langle\rangle} \to \Pi^{\langle\rangle}\downarrow$ when $n = 0$. Since terms of type $\langle\rangle$ are formulas, we may restate this case of theorem as $\varphi \to \varphi\downarrow$, i.e., truth implies existence.

Substitution of Identicals. To state the axiom for the substitution of identicals, let t be any type and let α and β be any variables of type t. Then the closures of the following are axioms:

(.9) $\alpha = \beta \to (\varphi \to \varphi')$, whenever β is substitutable for α in φ, and φ' is the result of replacing zero or more free occurrences of α in φ with occurrences of β.

$\star$**Actuality** (Fragile). We state the modally fragile axiom for the logic of actuality as before. We take only the universal closures of the following as axioms:

(.10) $\mathcal{A}\varphi \to \varphi$

In what follows, we refer to this axiom as (935.10)★, using our convention of marking modally fragile axioms with a ★. But all the other axioms of typed object theory will be referenced as (935.x), where $x \neq 10$.

Actuality (Necessary). We take *all* the closures of the following as axioms:

(.11) $\mathcal{A}\neg\varphi \equiv \neg\mathcal{A}\varphi$

(.12) $\mathcal{A}(\varphi \to \psi) \equiv (\mathcal{A}\varphi \to \mathcal{A}\psi)$

(.13) $\mathcal{A}\forall\alpha\varphi \equiv \forall\alpha\mathcal{A}\varphi$

(.14) $\mathcal{A}\varphi \equiv \mathcal{A}\mathcal{A}\varphi$

Necessity. The axioms for the necessity operator are also unchanged, with the exception of the last, which is restricted to the type for individuals. We take the closures of the following as axioms:

(.15) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$ (K)

(.16) $\Box\varphi \to \varphi$ (T)

(.17) $\Diamond\varphi \to \Box\Diamond\varphi$ (5)

(.18) $\Diamond\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$,
where x is a variable of type i and $E!$ is a constant of type $\langle i \rangle$

Necessity and Actuality. The axioms for necessity and actuality are also as before; we take the closures of the following as axioms:

(.19) $\mathcal{A}\varphi \to \Box\mathcal{A}\varphi$

(.20) $\Box\varphi \equiv \mathcal{A}\Box\varphi$

Descriptions. The axiom for definite descriptions governs descriptions of any type. Where t is any type and α and β are variables of type t, we assert the closures of the following as axioms:

(.21) $\alpha = \imath\alpha\varphi \;\equiv\; \forall\beta(\mathcal{A}\varphi^{\beta}_{\alpha} \equiv \beta = \alpha)$, provided β is substitutable for α in φ and doesn't occur free in φ

Relations. To state the axioms governing relations, let $t_1, \ldots, t_n$ be any types, $\alpha_1, \ldots, \alpha_n$ be distinct variables of types $t_1, \ldots, t_n$, respectively. Then we assert:

(.22) Where $O!$ has type $\langle\langle t_1, \ldots, t_n\rangle\rangle$, the closures of the following are axioms:

$$[\lambda\alpha_1 \ldots \alpha_n\, \varphi]\!\downarrow\; \to O![\lambda\alpha_1 \ldots \alpha_n\, \varphi] \qquad (n \geq 1)$$

Note: when $n=0$, $[\lambda\,\varphi]\!\downarrow\; \to O![\lambda\,\varphi]$ will be derivable from the next axiom.

(.23) Where $O!$ has type $\langle\langle\,\rangle\rangle$, the closures of the following are axioms:

$O!\varphi$, provided φ is not in $Base^{\langle\,\rangle}$, i.e., provided φ is not a constant of type $\langle\,\rangle$, a variable of type $\langle\,\rangle$, or a description of type $\langle\,\rangle$

(.24) Where $A!$ has type $\langle\langle t_1,\ldots,t_n\rangle\rangle$ and F a variable with type $\langle t_1,\ldots,t_n\rangle$, the closures of the following are axioms:

$$A!F \to \neg\exists\alpha_1\ldots\exists\alpha_n F\alpha_1\ldots\alpha_n \qquad (n \geq 0)$$

(.25) α-Conversion. Where $\alpha_1,\ldots,\alpha_n$ are any distinct variables and $[\lambda\alpha_1\ldots\alpha_n\varphi]'$ is any alphabetic variant of $[\lambda\alpha_1\ldots\alpha_n\varphi]$, the closures of the following are axioms:

$$[\lambda\alpha_1\ldots\alpha_n\,\varphi]\!\downarrow\; \to [\lambda\alpha_1\ldots\alpha_n\,\varphi] = [\lambda\alpha_1\ldots\alpha_n\,\varphi]' \qquad (n \geq 0)$$

(.26) β-Conversion. The closures of the following are axioms:

$$[\lambda\alpha_1\ldots\alpha_n\,\varphi]\!\downarrow\; \to ([\lambda\alpha_1\ldots\alpha_n\,\varphi]\alpha_1\ldots\alpha_n \equiv \varphi) \qquad (n \geq 0)$$

(.27) η-Conversion: Where F is any variable of type $\langle t_1,\ldots,t_n\rangle$ and $O!$ has type $\langle\langle t_1,\ldots,t_n\rangle\rangle$, the closures of the following are axioms governing elementary λ-expressions:

$$O!F \to ([\lambda\alpha_1\ldots\alpha_n\,F\alpha_1\ldots\alpha_n] = F) \qquad (n \geq 0)$$

So when $n = 0$, p is a variable of type $\langle\,\rangle$ and $O!$ has type $\langle\langle\,\rangle\rangle$, the closures $O!p \to ([\lambda\,p] = p)$ are axioms.

(.28) $([\lambda\alpha_1\ldots\alpha_n\,\varphi]\!\downarrow \;\&\; \Box\forall\alpha_1\ldots\forall\alpha_n(\varphi \equiv \psi)) \;\to\; [\lambda\alpha_1\ldots\alpha_n\,\psi]\!\downarrow \qquad (n \geq 1)$

A careful reading of Remark (931) should help one to understand the foregoing statements of axioms (.22) – (.28). But a few remarks may still prove useful, before we turn to the axioms of encoding.

(.22) asserts that significant n-ary λ-expressions ($n \geq 1$) denote *ordinary* relations. We'll later establish that $O![\lambda\,\varphi]$, where $O!$ has type $\langle\langle\,\rangle\rangle$, is a theorem, and this explains why the 0-ary case of (.22), i.e., $[\lambda\,\varphi]\!\downarrow \to O![\lambda\,\varphi]$, is derivable. (.23) asserts that a formula not among the basic expressions of type $\langle\,\rangle$ (i.e., not in $Base^{\langle\,\rangle}$) denotes an *ordinary* relation. (.24) asserts that *abstract* relations fail to be exemplified. (.25) and (.26), i.e., α- and β-Conversion, respectively, may be read exactly as (48.1) and (48.2), respectively, but have wider, type-theoretic significance — though β-Conversion (.26) now asserts instances when $n \geq 0$.[397]

[397] This differs from the second-order counterpart (48.2) in which $n \geq 1$. In second-order object theory, we were able to prove that $[\lambda\,\varphi] \equiv \varphi$, for any formula φ, as a theorem (111.2), and so the 0-ary case of β-Conversion, i.e., $[\lambda\,\varphi]\!\downarrow \to ([\lambda\,\varphi] \equiv \varphi)$ was derivable. But the proof of $[\lambda\,\varphi] \equiv \varphi$ rested on the proof of $[\lambda\,\varphi] = \varphi$, for any formula φ (111.2). But this latter schema, for reasons explained in Remark (931.1.c), can't be a theorem of typed-object theory. Since there is no longer an obvious means of deriving the 0-ary case of β-Conversion, it has to be included as an axiom.

(.27) asserts that η-Conversion holds only when the 'head' relation of an elementary λ-expression is *ordinary*.[398] Finally, (.28) may be read exactly as (49), but now has wider, type-theoretic significance.

Encoding. Let $t_1,\ldots,t_n$ be any types, $\alpha_1,\ldots,\alpha_n$ and $\beta_1,\ldots,\beta_n$ be any variables having types $t_1,\ldots,t_n$, respectively, G be a variable of type $\langle t_1,\ldots,t_n\rangle$, and $O!$ is of type $\langle\langle t_1,\ldots,t_n\rangle\rangle$. Then we assert the closures of following as axioms:

(.29) $\alpha_1\ldots\alpha_n G \equiv$
$\alpha_1[\lambda\beta_1\, G\beta_1\alpha_2\ldots\alpha_n]$ & $\alpha_2[\lambda\beta_2\, G\alpha_1\beta_2\alpha_3\ldots\alpha_n]$ & $\ldots$ & $\alpha_n[\lambda\beta_n\, G\alpha_1\ldots\alpha_{n-1}\beta_n]$ $(n\geq 2)$

Furthermore, let t be any type, α be any variable of type t, F be a variable of type $\langle t\rangle$, and $A!$ have type $\langle t\rangle$. Then we assert the closures of following as axioms:

(.30) $\alpha F\rightarrow\Box\alpha F$

(.31) $O!\alpha\rightarrow\neg\exists F\alpha F$

(.32) $\exists\alpha(A!\alpha\ \&\ \forall F(\alpha F\equiv\varphi))$, where φ has no free αs

No intuitive reading of these axioms is needed, as they are all type-theoretic counterparts of axioms in second-order object theory.

(**936**) **Remark:** How Typed Object Theory Generalizes Second-Order Object Theory. There are two ways in which the axioms of typed object theory generalize those of second-order object theory: (i) they reduce to those of second-order object theory under certain conditions, and (ii) when the theorems of second-order object theory are translated into typed object theory in a way that *preserves* exactly what they express in their original setting, they remain theorems. To see (i), consider the following conditions that would limit the language and axioms of typed object theory:

- restrict the types to i and to those having the form $\langle\underbrace{i,\ldots,i}_{n\text{ times}}\rangle$, where $n\geq 0$,
- eliminate from the language any expression that is, or contains, an expression whose type is not among the remaining types,
- eliminate from the language definite descriptions of relational type, i.e., descriptions of the form $\imath\alpha^t\varphi$, where $t\neq i$,

[398] The antecedent $O!F$ is a crucial condition, for without it, we could easily derive a contradiction. Suppose, for reductio, that (.27) asserted only $[\lambda\alpha_1\ldots\alpha_n\, F\alpha_1\ldots\alpha_n]=F$. Then consider the case in which F is (assigned) some abstract relation (as value). Then since the λ-expression satisfies the conditions of (.5), i.e., it is a core λ-expression because no variable bound by the λ is in encoding position in φ, the λ-expression would be significant. So by (.22), it would denote an ordinary relation. But then (.27.) would equate an ordinary relation with an abstract one. Contradiction.

- revert the definition of *proposition existence* (933.6.a) to its second-order counterpart (20.3), and

- eliminate all the unnecessary clauses from the definitions of *identity* for properties, relations, and propositions (933.9) – (933.12), that definitions become equivalent to their counterparts in second-order object theory.

Then, under these restrictions, it can be seen by inspection that:

- axioms (935.1) – (935.21) have only formulas of second-order object theory as instances,

- axioms (935.22) – (935.24) are inexpressible and so disappear, since there are no higher-order terms $A!$ and $O!$ having type $\langle t \rangle$ for $t \neq i$,

- axioms (935.25) and (935.26) reduce to their second-order counterparts (48.1) and (48.2) respectively, though the 0-ary case of (935.26) can be derived in second-order object theory and so isn't axiomatic in that system,

- the antecedent of (935.27) is inexpressible and eliminable, so that the axiom reduces to of (48.3),

- axiom (935.28) has only formulas of second-order object theory as instances, and

- (935.29) – (935.33) have only the formulas of second-order object theory as instances – all the higher-order instances become inexpressible.

It is also straightforward to see that (ii): when the theorems of second-order object theory are translated into typed object theory in a way that *preserves* exactly what they express in their original setting, they remain theorems. Here are two examples:

- The definition of proposition existence (20.3) is $p\!\downarrow =_{df} [\lambda x\, p]\!\downarrow$. This definition implies the theorem $p\!\downarrow \equiv [\lambda x\, p]\!\downarrow$. When we translate this theorem into typed object theory, we have to remember that the propositions of second-order object theory are all ordinary. So, when $p\!\downarrow \equiv [\lambda x\, p]\!\downarrow$ is translated into type-object theory in a way that preserves exactly what it expresses in second-order object theory, it becomes $O!p \rightarrow (p\!\downarrow \equiv [\lambda x\, p]\!\downarrow)$, where $O!$ has type $\langle\langle\rangle\rangle$, p has type $\langle\rangle$, and x has type i. Then, to see that this latter is a theorem, assume $O!p$, to show $p\!\downarrow \equiv [\lambda x p]\!\downarrow$. ($\rightarrow$) Assume $p\!\downarrow$. Now $\forall q([\lambda x q]\!\downarrow)$ is a closure of $[\lambda x q]\!\downarrow$ and so an axiom, by (935.5). Hence $[\lambda x\, p]\!\downarrow$. ($\leftarrow$) Assume $[\lambda x\, p]\!\downarrow$. But we know $O!p$ by assumption. Hence $\exists F\, Fp$, where F is a variable with type $\langle\langle\rangle\rangle$. So by definition (933.6.b), it follows that $p\!\downarrow$.

- The axiom of encoding (51) $xF \rightarrow \Box xF$. When translated into typed object theory, this becomes $O!F \rightarrow (xF \rightarrow \Box xF)$, where $O!$ has type $\langle\langle i \rangle\rangle$, F has type $\langle i \rangle$, and x has type i. To see that this latter is a theorem, assume $O!F$. But then $xF \rightarrow \Box xF$ is an instance of (935.30), when α is the variable x and has type i, and F has type $\langle i \rangle$.

We leave it as an exercise to prove that when the theorems of second-order object theory are translated in an exact way, so as to express in the typed setting precisely what they express in the second-order setting, the resulting typed versions remain theorems.

15.4 The Deductive System and Basic Theorems

15.4.1 The System, Negations, Conditionals, Quantification

Many of the metadefinitions for the system *typed PLM* are just analogous, if not identical, to their counterparts in second-order object theory.

(**937**) **Metadefinitions and Metarules:** The System of Typed PLM. The following metadefinitions and justifications (metatheorems) can be imported from second-order PLM by (a) referencing the axioms of typed PLM instead of the axioms of second-order PLM, (b) revising the metadefinitions and justifications of each second-order notion or metarule by substituting the new type-theoretic notions for second-order ones, and (c) decorating the terms in formulas with appropriate types if necessary:

(.1) Primitive Rule: Modus Ponens (58)

(.2) Derivations, Proofs, and Theoremhood (59)

(.3) Modally Strict: Derivations, Proofs, and Theoremhood (60)

(.4) Fundamental Properties of $\vdash$ and $\vdash_\Box$ (63)

(.5) Dependence (65)

(.6) Metarule GEN (66)

(.7) Metarule RN (68)

(.8) Primitive Metarule: Rule of Definition by Equivalence (72)

(.9) Primitive Metarule: Rule of Definition by Identity (73)

In what follows, we'll refer to these, respectively, as (937.1) [58], (937.2) [59], etc. The proofs of (937.4) [63], (937.6) [66], and (937.7) [68] are left as exercises.

Note that in the statement of (.9), the general form of definitions by equivalence should cite the appropriate types, so that in a definition of the general form $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$, the variables $\alpha_1,\ldots,\alpha_n$ may have any types $t_1,\ldots,t_n$, respectively, and in any instance of the definition of the form $\tau(\tau_1,\ldots,\tau_n) =_{df} \sigma(\tau_1,\ldots,\tau_n)$, the terms $\tau_1,\ldots,\tau_n$ should be substitutable for $\alpha_1,\ldots,$ α_n, respectively, in $\sigma(\alpha_1,\ldots,\alpha_n)$ (which ensures that $\tau_1,\ldots,\tau_n$ have types $t_1,\ldots,t_n$, respectively).

Finally, we henceforth use the Conventions Regarding Metarules as stated in (67) but with the reference to axiom (43)★ in that statement replaced by its type-theoretic counterpart axiom (935.10)★. These conventions allow us to state only metarules for $\vdash$, without having to repeat a version of the rule for $\vdash_\Box$, unless the conditions of the metarule specifically require the existence of a modally strict derivation or proof (i.e., unless the conditions of the metarule specifically appeal to some claim of the for $\Gamma \vdash_\Box \varphi$ or $\vdash_\Box \varphi$). Moreover, we continue to disallow metarules whose justification or conditions of application depend on (935.10)★ or any other modally fragile axiom.

(**938**) **Theorems and Metarules/Derived Rules:** Propositional Logic.

> The theorem schemata and metarules governing negations and conditionals apply in complete generality to the language of typed PLM.
>
> (74) – (91)

We'll reference these schemata and rules in what follows by both the present and original item number. So, for example, the theorem $\varphi \to \varphi$ (74) will henceforth be referenced as (938) [74].

(**939**) **Theorems and Metarules/Derived Rules:** Quantificational Logic.

> The theorem schemata and metarules governing the quantifiers apply in complete generality to the language of typed PLM, provided the terms in each schema or metarule are appropriately typed.
>
> (93) – (103)

To take a simple example, consider the first form of Rule ∀E, which we'll henceforth reference as (939) [93.1]. The rule states: If $\Gamma_1 \vdash \forall\alpha\varphi$ and $\Gamma_2 \vdash \tau\downarrow$, then $\Gamma_1,\Gamma_2 \vdash \varphi^{\tau}_{\alpha}$, provided τ is substitutable for α in φ. This is now applies only when both τ and α both have type t, where t is any type. For only then will τ be substitutable for α.

15.4.2 Logical Existence, Identity, and Uniqueness

We now examine some of the main theorems and metatheorems articulated in Section 9.7 and indicate how they are to be formulated given the increased expressive power of typed object theory. The first change in the procession of theorems concerns the fact that (104.1) and (104.2) no longer hold for every formula φ. We start our discussion there.

(**940**) **Theorems:** Existence and Identity. It is axiomatic, by (935.5), that $[\lambda\,\varphi]\downarrow$, for any formula φ. But for the reasons given in Remark (931.1.a), (104.1) and (104.2) can't be theorems of typed object theory. We've now undermined these theorems by redefining proposition existence. Using the new definition of proposition existence (933.6.b), our first theorem becomes:

(.1) $\varphi\downarrow$, where φ is any formula other than a description of type $\langle\,\rangle$

Though this theorem doesn't assert that descriptions of type $\langle\,\rangle$ are significant, there is one group of descriptions of this type that are provably significant, though we're not yet in a position to show this. Since the definition of a *canonical description* (253) extends in type theory to allow for canonical descriptions of every type, we'll later extend (.1) to include canonical descriptions of type $\langle\,\rangle$, which are formulas by the BNF. Indeed, the comprehension and identity principles for abstract objects, along with our axiom for descriptions, will guarantee that canonical descriptions of *every* type t are significant. So canonical descriptions of type $\langle\,\rangle$ will be significant.

Note that the BNF also tells us that the formulas constitute all the terms of type $\langle\,\rangle$. So where Π is any term of type $\langle\,\rangle$, it follows from (.1) that (.2) Π exists, provided Π is not a description:

(.2) $\Pi\downarrow$, where Π is any term of type $\langle\,\rangle$ other than a description

Again, this will be extended later, since canonical descriptions of type $\langle\,\rangle$ will also be provably significant.

Note also that (.1) is the type-theoretic version of (104.2), while (.2) is the type-theoretic version of (104.1). The difference in the order of presentation isn't portentous, but rather reflects only the fact that the change in the definition of proposition existence alters the way in which the typed versions are proved. In typed object theory, (.1) is proved by cases and each of the two cases holds axiomatically. Then (.2) follows from (.1).

(**941**) **Remark:** Fact About Dependencies and Generality in the Move from Second-Order to Typed Object Theory. Since the theorems in (104) do not transfer to typed object theory in completely general form, it becomes important to consider whether the theorems that depend on (104) also fail to transfer to typed object theory in general form. It should be observed, however, that from the fact that:

- a second-order theorem χ doesn't transfer to typed object theory in complete generality,

it doesn't follow that:

- the type-theoretic version of every second-order theorem θ that depends on χ fails to be provable in typed object theory in complete generality.

It may be that θ can be proved in complete generality by other means. We'll some nice examples of this in what follows:

- The second-order theorem (107.1), i.e., $\tau = \sigma \to \tau\downarrow$ (for any terms τ and σ of the same type), was derived from theorem (104.1), namely that $\Pi\downarrow$ holds for any 0-ary relation term Π^0. But though the type-theoretic version of (104.1), namely (940.2), doesn't hold in complete generality, the type-theoretic version of (107.1) does hold in complete generality. In (943.1) below, we prove that $\tau = \sigma \to \tau\downarrow$ holds for any terms τ and σ of type t, for any t.
- The second-order theorem (111.5), i.e., that $\varphi \equiv \varphi'$ (for any alphabetically variant formulas φ and φ'), was derived from theorem (111.4), namely, that $\varphi = \varphi'$ for any alphabetic variants φ and φ'. But the type-theoretic versions of (111.4) are restricted – one can prove only that $\varphi\downarrow \to (\varphi = \varphi')$ (950.5) and that $\varphi = \varphi'$ whenever φ is any formula other than a description of type $\langle\rangle$ (950.6). Nevertheless, it is provable in complete generality that $\varphi \equiv \varphi'$, for any alphabetically variant formulas φ and φ' (950.7).

(**942**) **Theorems:** Logical Existence is Necessary. Observe next that existence statements are necessary if true, so that (106) holds in complete generality for typed object theory, i.e.,

$$\tau\downarrow \to \Box\tau\downarrow$$

The proof is analogous to that of (106).

(**943**) **Theorems:** Identity Implies Existence. True identity statements imply the significance of the terms flanking the identity sign, so that (107.1) and (107.2) also hold in complete generality. But the proofs of the type-theoretic versions differ somewhat from those of the second-order versions:

(.1) $\tau = \sigma \to \tau\downarrow$

(.2) $\tau = \sigma \to \sigma\downarrow$

(**944**) **Theorems:** Identical Relations are Necessarily Equivalent. Type-theoretic versions of (108.1) and (108.2) now easily follow. (.1) If an identity holds between Π and Π', then the relations they signify are necessarily equivalent:

(.1) $\Pi = \Pi' \to \Box\forall x_1 \dots \forall x_n(\Pi x_1 \dots x_n \equiv \Pi' x_1 \dots x_n)$, provided Π and Π' are terms of type $\langle t_1, \dots, t_n\rangle$ ($n \geq 0$) and $x_1, \dots, x_n$ are variables of type $t_1, \dots, t_n$, respectively, that don't occur free in Π and Π'.

The converse, however, does not hold.

When $n=0$ and Π and Π' are terms of type $\langle\rangle$ and so formulas, an instance of the above asserts that (.2) if φ is identical to ψ, then necessarily, φ if and only ψ:

(.2) $\varphi = \psi \rightarrow \Box(\varphi \equiv \psi)$

(**945**) **Metarules/Derived Rules:** Substitution of Identicals. The foregoing facts ensure that the Rule for the Substitution of Identicals (110), in its type-theoretic guise, is still justified:

> **Rule =E**
> Let t be any type, τ and σ be any terms of type t, and α be any variable of type t. Then if $\Gamma_1 \vdash \varphi^{\tau}_{\alpha}$ and $\Gamma_2 \vdash \tau = \sigma$, then $\Gamma_1, \Gamma_2 \vdash \varphi'$, whenever τ and σ are any terms substitutable for α in φ, and φ' is the result of replacing zero or more occurrences of τ in φ^{τ}_{α} with occurrences of σ.
> [Variant: φ^{τ}_{α}, $\tau = \sigma \vdash \varphi'$]

In the usual way, the Variant is a derived rule.

(**946**) **Theorems:** Objects are Ordinary or Abstract. We now show that (.1) *being ordinary*$^{\langle t\rangle}$ exists, for any type t; (.2) *being abstract*$^{\langle t\rangle}$ exists, for any type t; and (.3) every object of type t is either ordinary or abstract:

(.1) $O!\downarrow$, where $O!$ has type $\langle t\rangle$, for any type t

(.2) $A!\downarrow$, where $A!$ has type $\langle t\rangle$, for any type t

(.3) $O!x \vee A!x$, where x is a variable of type t, and $O!$ and $A!$ have type $\langle t\rangle$

(**947**) **Theorems:** Identity is an Equivalence Condition. Where t is any type, and x, y, and z are variables of type t, we have (.1) x is identical to x, (.2) if x is identical to y, then y is identical to x, and (.3) if x is identical to y and y is identical to z, then x is identical to z:

(.1) $x = x$

(.2) $x = y \rightarrow y = x$

(.3) $x = y \;\&\; y = z \rightarrow x = z$

It also follows that (.4) if x is identical to y, then for any z, x is identical to z if and only if y is identical to z:

(.4) $x = y \equiv \forall z(x = z \equiv y = z)$

(**948**) **Metarules/Derived Rules:** Rule of Identity Introduction. Rule =I (118) holds in typed object theory:

(.1) **Rule =I**
Where τ is any term of any type:
If $\Gamma \vdash \tau\downarrow$, then $\Gamma \vdash \tau = \tau$ [Variant: $\tau\downarrow \vdash \tau = \tau$]

(.2) **Rule =I** (Special Case)
$\vdash \tau = \tau$, provided τ is a primitive constant, a variable of any type, or a core λ-expression.

(**949**) **Metarules/Derived Rules:** Identity by Definition, $=_{df}$E, and $=_{df}$I. Let $\sigma(\tau_1,\ldots,\tau_n)$ and $\tau(\tau_1,\ldots,\tau_n)$ abbreviate $\sigma^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$ and $\tau^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$, respectively ($1 \leq i \leq n$). Then it is straightforward to justify the following:

(.1) **Rule of Identity by Definition**
Let $t_1,\ldots,t_n$ be any types ($n \geq 0$) and let $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$ be a definition-by-= in which the variables $\alpha_1,\ldots,\alpha_n$ ($n \geq 0$) occur free and have types $t_1,\ldots,t_n$, respectively. Let $\tau_1,\ldots,\tau_n$ be terms of types $t_1,\ldots,t_n$, respectively, that are substitutable, respectively, for $\alpha_1,\ldots,\alpha_n$ in both the definiens and definiendum. Then:

$$\text{if } \Gamma \vdash \sigma(\tau_1,\ldots,\tau_n)\downarrow, \text{ then } \Gamma \vdash \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n)$$

We immediately obtain the classical introduction and elimination rules for the definiendum:

(.2.a) **Rule of Definiendum Elimination:** (Rule $=_{df}$**E**)
Let $t_1,\ldots,t_n$ be any types ($n \geq 0$) and let $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$ be a definition-by-= in which the variables $\alpha_1,\ldots,\alpha_n$ ($n \geq 0$) occur free and have types $t_1,\ldots,t_n$, respectively. Let $\tau_1,\ldots,\tau_n$ be terms of types $t_1,\ldots,t_n$, respectively, and substitutable for $\alpha_1,\ldots,\alpha_n$, respectively, in both the definiens and definiendum. Furthermore, let φ be any formula that contains one or more occurrences of $\tau(\tau_1,\ldots,\tau_n)$ and let φ' be the result of replacing zero or more occurrences of $\tau(\tau_1,\ldots,\tau_n)$ in φ by $\sigma(\tau_1,\ldots,\tau_n)$. Then if $\Gamma \vdash \sigma(\tau_1,\ldots,\tau_n)\downarrow$ and $\Gamma \vdash \varphi$, then $\Gamma \vdash \varphi'$.
[Variant: $\sigma(\tau_1,\ldots,\tau_n)\downarrow, \varphi \vdash \varphi'$]

(.2.b) **Rule of Definiendum Introduction:** (Rule $=_{df}$**I**)
Let $t_1,\ldots,t_n$ be any types ($n \geq 0$) and let $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$ be a definition-by-= in which the variables $\alpha_1,\ldots,\alpha_n$ ($n \geq 0$) occur free and have types $t_1,\ldots,t_n$, respectively. Let $\tau_1,\ldots,\tau_n$ be terms of types $t_1,\ldots,t_n$, respectively, and substitutable for $\alpha_1,\ldots,\alpha_n$, respectively, in both the definiens and definiendum. Furthermore, let φ be any formula that contains one or more occurrences of $\sigma(\tau_1,\ldots,\tau_n)$ and let φ' be the result of replacing zero or more occurrences of $\sigma(\tau_1,\ldots,\tau_n)$ in φ by $\tau(\tau_1,\ldots,\tau_n)$. Then if $\Gamma \vdash \sigma(\tau_1,\ldots,\tau_n)\downarrow$ and $\Gamma \vdash \varphi$, then $\Gamma \vdash \varphi'$.
[Variant: $\sigma(\tau_1,\ldots,\tau_n)\downarrow, \varphi \vdash \varphi'$]

(**950**) **Theorems:** Identity and Alphabetic Variants of Formulas and Proposition Terms. We next come to another departure from the theorems in Chapter 9, concerning (111.1). As we saw in (931.1.c), $[\lambda\,\varphi] = \varphi$, for arbitrary φ, cannot be a theorem. But the following are theorems:

(.1) $O!\varphi \rightarrow ([\lambda\,\varphi] = \varphi)$, where φ is any formula

(.2) $[\lambda\,\varphi] = \varphi$, where φ is any non-basic formula, i.e., any formula that is not in $Base^{\langle\rangle}$

Our next theorem is an equivalence. Recall that in second-order object theory, the axiom β-Conversion (48.2) was stated only for $n \geq 1$. We didn't need to assert the 0-ary case, i.e., $[\lambda\,\varphi]\!\downarrow \rightarrow ([\lambda\,\varphi] \equiv \varphi)$, as an axiom because $[\lambda\,\varphi] \equiv \varphi$ was derivable as a theorem (111.2) and so the 0-ary case of β-Conversion followed trivially (by the truth of its consequent). But we used theorem (111.1), i.e., $[\lambda\,\varphi] = \varphi$, to derive (111.2). Although $[\lambda\,\varphi] = \varphi$ doesn't hold in full generality in typed object theory, we can still prove $[\lambda\,\varphi] \equiv \varphi$ because the type-theoretic version of the axiom β-Conversion (935.26) has been strengthened so that it holds for $n = 0$. Thus, $[\lambda\,\varphi]\!\downarrow \rightarrow ([\lambda\,\varphi] \equiv \varphi)$ is axiomatic and so it is an immediate consequence of this axiom and axiom (935.5) that:

(.3) $[\lambda\,\varphi] \equiv \varphi$

The *theory of truth* therefore remains intact in typed object theory, since the above asserts: *that*-φ is true if and only if φ. It should be remembered, though, that $[\lambda\,\varphi]\!\downarrow\, \equiv \varphi\!\downarrow$ is not a theorem. As we saw in (931.1.d.iii), the formula $[\lambda\,\imath p(p \,\&\, \neg p)]\!\downarrow\, \equiv \imath p(p \,\&\, \neg p)\!\downarrow$ is provably false; the left condition is axiomatically true (935.5), but the right condition provably fails to be true.

Now (111.3), i.e., the claim $[\lambda\,\varphi] = [\lambda\,\varphi]'$, for alphabetic variants $[\lambda\,\varphi]$ and $[\lambda\,\varphi]'$, remains a theorem of object theory, in complete generality:

(.4) $[\lambda\,\varphi] = [\lambda\,\varphi]'$, where $[\lambda\,\varphi]$ and $[\lambda\,\varphi]'$ are alphabetic variants

We turn next to alphabetic variants of formulas other that those of the form $[\lambda\,\psi]$. In general, it follows that (.5) if φ is a significant formula, then an identity holds between φ and any of its alphabetic variants φ':

(.5) $\varphi\!\downarrow \rightarrow (\varphi = \varphi')$, where φ' is any alphabetic variant of φ

Note that this immediately implies $O!\varphi \rightarrow \varphi = \varphi'$, since $O!\varphi$ implies $\varphi\!\downarrow$. It is also an immediate consequence of (.5) and (940.1) that:

(.6) $\varphi = \varphi'$, where φ is any formula other than a description of type $\langle\rangle$ and φ' is any alphabetic variable of φ

Clearly, (.6) must exclude descriptions generally, since $\imath p(p \,\&\, \neg p) = \imath q(q \,\&\, \neg q)$ should not be a theorem (true identity statements imply the significance of the terms flanking the identity sign and the descriptions flanking the identity sign here fail to be significant, on pain of contradiction). But, later, once we establish that canonical descriptions of every type t are significant, it will follow that *canonical* descriptions of type $\langle\,\rangle$ are significant and since the latter are formulas, it will follow from (.6) that an identity holds between alphabetically-variant *canonical* descriptions. More generally, an identity holds between any *significant* description and any of its alphabetic variants, i.e., it will be a theorem that $\imath\alpha^t\varphi\!\downarrow \to (\imath\alpha^t\varphi = (\imath\alpha^t\varphi)')$, where $(\imath\alpha^t\varphi)'$ is any alphabetic variant of $\imath\alpha^t\varphi$, for any type t (957) [154].

If we now consider equivalences instead of identities, then it is a theorem that the *equivalence* of alphabetically-variant formulas holds unconditionally:

(.7) $\varphi \equiv \varphi'$

Clearly, this holds even for formulas that aren't significant. For example, the biconditional $\imath p(p \,\&\, \neg p) \equiv \imath q(q \,\&\, \neg q)$ holds because both sides are (necessarily) false. Finally, we have:

(.8) $(\varphi \equiv \psi) \equiv ([\lambda\, \varphi] \equiv [\lambda\, \psi])$

(**951**) **Metarule/Derived Rule:** Rule of Alphabetic Variants. The classical Rule of Alphabetic Variants holds in typed object theory:

Rule of Alphabetic Variants
$\Gamma \vdash \varphi$ if and only if $\Gamma \vdash \varphi'$, where φ' is any alphabetic variant of φ
[Variant $\varphi \dashv\vdash \varphi'$]

As a special case, when $\Gamma = \varnothing$, our rule asserts that a formula is a theorem if and only if all of its alphabetic variants are theorems.

Now that we've established that alphabetically variant formulas are equivalent and the associated Rule of Alphabetic Variants, we can start expressing quantified claims by using object language variables $\forall x\varphi$, $\exists y\psi$, etc. (where x and y are of any type), instead of as $\forall\alpha\varphi$, $\exists\beta\psi$, etc. (where α and β are of any type). Heretofore, we had used the metavariables α, β, etc., to assert axioms and theorems to be assured that the claims hold for any appropriate variables of the same type, to be assured that the claims hold for all alphabetic variants. But with the Rule of Alphabetic Variants, we can revert to using more readable object language variables.

(**952**) **Theorems:** Negative Free Logic. The axioms of negative free logic, suitably typed, are now derivable:

(.1) $\tau\!\downarrow \equiv \exists x(x = \tau)$, provided that x is a variable having the same type as τ and x doesn't occur free in τ

(.2) $\forall x\varphi \to (\exists y(y=\tau) \to \varphi^{\tau}_{x})$, provided that (a) x and y are variables having the same type as τ, (b) τ is substitutable for x in φ, and (c) y doesn't occur free in τ

(.3) $\exists x(x=\tau)$, provided (a) τ is either a primitive constant, a variable, or a core λ-expression, and (b) x is a variable of the same type as τ that doesn't occur free in τ

(.4) $(\Pi^n\kappa_1\ldots\kappa_n \vee \kappa_1\ldots\kappa_n\Pi^n) \to \exists\beta(\beta=\tau)$, where τ is any of Π^n, $\kappa_1, \ldots,$ or κ_n, and β is a variable of the same type as τ that doesn't occur free in τ.

The proofs follow the corresponding theorems of second-order object theory.

(**953**) **Theorems:** Necessarily Everything of Type t Necessarily Exists. It is sufficient to reprise just one pair of theorems from (123) in type-theoretic form. Where t is any type and x and y are distinct variables of type t, it is a consequence of our axioms and rules that:

(.1) $\Box\forall x\Box x{\downarrow}$

(.2) $\Box\forall x\Box\exists y(y=x)$

(**954**) **Theorems:** Identity and Necessity. It is a consequence of the foregoing that for any type t, (.1) necessarily everything of type t is self-identical; (.2) everything of type t is necessarily self-identical; and (.3) things of type t are identical if and only if they are necessarily identical:

(.1) $\Box\forall x(x=x)$, for any variable x of any type t

(.2) $\forall x\Box(x=x)$, for any variable x of any type t

(.3) $\tau=\sigma \equiv \Box\tau=\sigma$, where τ and σ are any terms of type t, for any type t

(**955**) **Theorems and Definitions:** Quantification, Identity, Uniqueness, and Necessity. The type-theoretic versions of the theorems and definitions in items (126) – (129) are now all derivable.

- Theorems of Quantification and Identity (126):

  (.1) $\varphi \equiv \exists y(y{=}x \,\&\, \varphi^{y}_{x})$, provided y is substitutable for x in φ and doesn't occur free in φ, for any variables x and y of any type t

  (.2) $\tau{\downarrow} \to (\varphi^{\tau}_{x} \equiv \exists x(x=\tau \,\&\, \varphi))$, provided τ is substitutable for x in φ, for any variable x and term τ of type t, for any type t

  (.3) $(\varphi \,\&\, \forall y(\varphi^{y}_{x} \to y=x)) \equiv \forall y(\varphi^{y}_{x} \equiv y=x)$,
  provided x and y are distinct variables of type t, for any type t, and y is substitutable for x in φ and doesn't occur free in φ

(.4) $(\varphi^y_x \,\&\, \forall x(\varphi \to x = y)) \equiv \forall x(\varphi \equiv x = y)$,
provided x and y are distinct variables of any type t, and y is substitutable for x in φ and doesn't occur free in φ.

- Alternative Definitions of Uniqueness in (127)

(.5) $\exists!x\varphi \;\equiv_{df}\; \exists x(\varphi \,\&\, \forall y(\varphi^y_x \to x = y))$,
provided y doesn't occur free, and is substitutable for x, in φ, for any variables x and y of any type t

(.6) $\exists!x\varphi \;\equiv_{df}\; \exists x \forall y(\varphi^y_x \equiv y = x)$,
provided y doesn't occur free, and is substitutable for x, in φ, for any variables x and y of any type t

- Uniqueness Implies At Most One (128):

(.7) $\exists!x\varphi \to \forall y \forall z((\varphi^y_x \,\&\, \varphi^z_x) \to y = z)$,
provided y and z don't occur free, and are substitutable for x, in φ, for any variables x, y, and z of any type t

- Uniqueness and Necessity (129):

(.8) $\forall x(\varphi \to \Box\varphi) \to (\exists!x\varphi \to \exists!x\Box\varphi)$, for any variable x of any type t

15.4.3 The Theory of Actuality and Descriptions

(**956**) **Theorems:** Basic Theorems of Actuality Preserved. The theorems and metarules for the actuality operator are preserved intact from second-order object theory:

(.1) The type-theoretic versions of theorems (131) – (134.4) are all straightforward and derivable; we henceforth refer to them as (956.1) [131] – (956.1) [134.4], respectively. Note that (956.1) [134.2] and (956.1) [134.3] can be written as follows, respectively:

- $\forall x \mathcal{A}(\mathcal{A}\varphi \to \varphi)$, where x is a variable of any type
- $\mathcal{A}\forall x(\mathcal{A}\varphi \to \varphi)$, where x is a variable of any type

and that (956.1) [134.4] can be written as:

- $\mathcal{A}\forall x_1 \ldots \forall x_n(\mathcal{A}\varphi \to \varphi)$ where $x_1, \ldots, x_n$ are variables of any types.

(.2) The type-theoretic version of the Rule of Actualization (135) is justified; we henceforth refer to it as Rule RA (956.2) [135].

(.3) The type-theoretic versions of the theorems governing actuality and negation, (138.1)★ and (138.2)★, are derivable; we henceforth refer to them as (956.3) [138.1]★ and (956.3) [138.2]★, respectively.

(.4) The type-theoretic versions of the modally strict theorems of actuality (139.1) – (139.11) are derivable; we henceforth refer to them as (956.4) [139.1] – (956.4) [139.11], respectively. Note that (956.4) [139.10] and (956.4) [139.11] can be written, respectively, as follows:

- $\mathcal{A}\exists x\varphi \equiv \exists x\mathcal{A}\varphi$, where x is a variable of any type
- $\mathcal{A}\forall x(\varphi \equiv \psi) \equiv \forall x(\mathcal{A}\varphi \equiv \mathcal{A}\psi)$, where x is a variable of any type

(.5) The non-modally strict equivalence of the conditions 'x is uniquely an actual φ' and 'x is uniquely φ' (140)★ holds in type theory:

- $\forall y(\mathcal{A}\varphi^y_x \equiv y\!=\!x) \equiv \forall y(\varphi^y_x \equiv y\!=\!x)$, provided x and y are both variables of type t, for some type t, and y is substitutable for x in φ and doesn't occur free in φ

We henceforth refer to this as (956.5) [140]★.

(**957**) **Theorems:** Typed Description Theory. Theorems (141)★ – (154) governing definite descriptions, suitably typed in the obvious way, all hold in complete generality in typed object theory. We henceforth reference these as (957) [141★] – (957) [154].

(**958**) **Theorems:** Some Facts About (Some Empty) Descriptions of the Empty Type. The following theorems are about a particular description of type $\langle\,\rangle$ that is empty. Consequently, F in (.1) has type $\langle\langle\,\rangle\rangle$, and $\downarrow$ is defined for propositions as in (933.6.b):

(.1) $\neg\exists F(F\imath p(p \,\&\, \neg p))$

(.2) $\neg(\imath p(p \,\&\, \neg p)\downarrow)$

(.3) $\neg\imath p(p \,\&\, \neg p)$

(.4) $\imath p\varphi \equiv \exists p(\mathcal{A}\varphi \,\&\, \forall q(\mathcal{A}\varphi^q_p \rightarrow q\!=\!p) \,\&\, p)$

In case there is any doubt, (.3) asserts: it is not the case that the proposition that is both true and not true is true. And (.4) asserts: The proposition p such that φ is true if and only if there is a proposition p such that (a) p is actually such that φ, and (b) any proposition q that is actually such that φ is identical to p. Note that $\imath p(p \,\&\, \neg p)$ is not a counterexample to (.4) – both sides are false with respect to this description. Results analogous to (.1) – (.4) hold for any other description of type $\langle\,\rangle$ that is provably empty.

15.4.4 The Theory of Necessity

(**959**) **Theorems and Metarules:** Quantified S5 Modal Logic Governs Typed Object Theory. The metarules and theorems derivable from the first principles governing quantified S5 modal logic, suitable typed, all hold in complete generality in typed object theory:

(.1) The metarules (157.1) – (157.4), i.e., RM, RM$\diamond$, RE, and RE$\diamond$, are all justifiable in typed object theory. We henceforth refer to these as (959.1) [157.1] – (959.1) [157.4], respectively.

(.2) The basic theorems of K in (158.1) – (158.16) all hold in typed object theory. We henceforth refer to them as (959.2) [158.1] – (959.2) [158.16], respectively.

(.3) Suitably typed, the metarules of Necessary Equivalence (159.1) – (159.4), and the Rules of Substitution (160.1) – (160.3), all hold in complete generality in typed object theory. We henceforth refer to them as (959.3) [159.1] – (959.3)[159.4], and (959.3) [160.1] – (959.3) [160.3], respectively.

(.4) The following modal theorems:

- (162.1) – (162.7) Additional K Theorems
- (163.1) – (163.2) T$\diamond$, 5$\diamond$ Schemata
- (164.1) – (164.5) Actuality, Negation, and Possibility
- (165.1) – (165.13) Basic S5 Theorems

all hold in complete generality in typed object theory. We henceforth refer to these, respectively, as:

- (959.4) [162.1] – (959.4) [162.7]
- (959.4) [163.1] – (959.4) [163.2]
- (959.4) [164.1] – (959.4) [164.5]
- (959.4) [165.1] – (959.4) [165.13]

(.5) The metarules resulting from the B and B$\diamond$ schemata, (166.1) – (166.2) all hold in complete generality in typed object theory. We henceforth refer to these as (959.5) [166.1] – (959.5) [166.2].

(.6) Suitably typed, the laws of quantified S5 modal logic, (167.1) – (167.4) and (168.1) – (168.6), all hold in complete generality in typed object theory. We henceforth refer to these, respectively, as (959.6) [167.1] – (959.6) [167.4] and (959.6) [168.1] – (959.6) [168.6]

(.7) The laws of modal collapse (169) – (180), suitably typed, all hold in typed object theory. We henceforth refer to these, respectively, as (959.7) [169] – (959.7) [180].

15.4.5 The Typed Theory of Relations

(**960**) **Theorems and Definitions:** Principles Governing Relation Terms Hold in Typed Object Theory.

(.1) The type-theoretic versions of Strengthened β-Conversion (181), its corollaries (183.1) – (183.2), and the Rules $\overrightarrow{\beta}$C (184.1) and $\overleftarrow{\beta}$C (184.2), all hold. We refer to these henceforth as (960.1) [183.1] – (960.1) [183.2], (960.1) [184.1], and (960.1) [184.2], respectively.

(.2) When ρ and Π^n are relation terms, each having some type $t \neq i$, the definitions:

(a) ρ is *elementary* (185.1)

(b) ρ is an *η-expansion of* Π^n (185.2)

(c) ρ is an *η-contraction of* Π^n (185.3)

(d) ρ' is an *immediate η-variant of ρ with respect to* Π^n (185.4)

(e) ρ' is an *η-variant of* ρ (185.5)

are all straightforwardly reformulable in typed object theory.

(.3) The η-Conversion lemmas (186.1) and (186.2), and the Rule ηC of η-Conversion (187) hold in typed object theory. We henceforth refer to these as (960.3) [186.1], (960.3) [186.2], and (960.3) [187], respectively.

(.4) The theorems that govern relation terms that differ by co-denoting descriptions, (188.1) and (188.2), both hold in typed object theory. We refer to these as (960.4) [188.1] and (960.4) [188.2], respectively.

Let's take as an example one of the principles in (.1). The type-theoretic version of Strengthened β-Conversion would be formulated as:

Where $t_1,\ldots,t_n$ are any types and $\alpha_1,\ldots,\alpha_n$ and $\beta_1,\ldots,\beta_n$ are, respectively, any variables of types $t_1,\ldots,t_n$, then it is a theorem that:

$$[\lambda\alpha_1 \ldots \alpha_n\, \varphi]\downarrow \to ([\lambda\alpha_1 \ldots \alpha_n \varphi]\beta_1 \ldots \beta_n \equiv \varphi^{\beta_1,\ldots,\beta_n}_{\alpha_1,\ldots,\alpha_n}),$$

provided $\beta_1,\ldots,\beta_n$ are substitutable for $\alpha_1,\ldots,\alpha_n$, respectively, in φ

And analogous type-theoretic formulations hold for other theorems noted in (.1), (.3) and (.4).

(**961**) **Theorems:** Facts About Relations. Our theory of identity for relations implies, for any type t: (.1) properties having type $\langle t\rangle$ that are encoded by the same objects of type t are necessarily encoded by the same objects of type t; (.2) ordinary properties having type $\langle t\rangle$ are identical whenever they are encoded by the same objects of type t; (.3) abstract properties having type $\langle t\rangle$ that encode the same properties having type $\langle\langle t\rangle\rangle$, necessarily encode those properties; and (.4) abstract properties having type $\langle t\rangle$ are identical whenever they encode the same properties having type $\langle\langle t\rangle\rangle$. We may represent these as follows, where t is any type, x is a variable of type t, $O!$ and $A!$ are defined constants of type $\langle t\rangle$, F and G are variables of type $\langle t\rangle$, and $\mathcal{H}$ is a variable of type $\langle\langle t\rangle\rangle$:

(.1) $\forall x(xF \equiv xG) \rightarrow \Box\forall x(xF \equiv xG)$

(.2) $(O!F \,\&\, O!G) \rightarrow (\forall x(xF \equiv xG) \rightarrow F = G)$

(.3) $\forall\mathcal{H}(F\mathcal{H} \equiv G\mathcal{H}) \rightarrow \Box\forall\mathcal{H}(F\mathcal{H} \equiv G\mathcal{H})$

(.4) $(A!F \,\&\, A!G) \rightarrow (\forall\mathcal{H}(F\mathcal{H} \equiv G\mathcal{H}) \rightarrow F = G)$

Note that (.2) and (.4) offer an *extensional* theory of properties and relations despite their hyperintensional character with respect to exemplification. (.2) easily yields $(O!F \,\&\, O!G) \rightarrow (F = G \equiv \forall x(xF \equiv xG))$. This intuitively tells us that ordinary properties are identical precisely when they have the same encoding extension. And (.4) easily yields $(A!F \,\&\, A!G) \rightarrow (F = G \equiv \forall\mathcal{H}(F\mathcal{H} \equiv G\mathcal{H}))$. This intuitively tells us that abstract properties are identical precisely when they are in the encoding extensions of the same properties of properties.

(.5) The Comprehension Principle for Ordinary n-ary Relations ($n \geq 1$) is a theorem schema asserting the existence of *ordinary* relations of higher type. Where $t_1,\ldots,t_n$ are any types ($n \geq 1$), F is a variable of type $\langle t_1,\ldots,t_n\rangle$, $O!$ is a property having type $\langle\langle t_1,\ldots,t_n\rangle\rangle$, $x_1,\ldots,x_n$ are variables having types $t_1,\ldots,t_n$, respectively, and φ is any formula such that (a) F doesn't occur free in φ and (b) $x_1,\ldots,x_n$ don't occur free in encoding position in φ, it is a theorem that:

$$\exists F(O!F \,\&\, \Box\forall x_1\ldots\forall x_n(Fx_1\ldots x_n \equiv \varphi))$$

When $n = 1$, this becomes the Comprehension Principle for Ordinary Properties.

(.6) Paradoxical λ-expressions aren't significant. Suitably typed versions of the theorems in (192) and (193)★ are theorems of typed object theory. We henceforth refer to these respectively as (961.6) [192.1] – (961.6) [192.5] and (961.6) [193.1]★ – (961.6) [193.2]★.

(.7) The Comprehension Principle for Propositions is a theorem schema governing *ordinary* propositions. Where p is a variable of type $\langle\,\rangle$, $O!$ is a property having type $\langle\langle\,\rangle\rangle$, and φ is any formula in which p doesn't occur free, it is a theorem that:

$$\exists p(O!p \,\&\, \Box(p \equiv \varphi))$$

(.8) The conditions implying distinctness of relations in (195) all hold, when appropriately typed, no matter whether F and G are (i) both ordinary relations, (ii) both abstract relations, or (iii) one is ordinary while the other is abstract. In what follows, we refer to these as (961.8) [195.1] – (961.8) [195.4].

Note that when F and G are abstract relations of some type $\langle t_1,\ldots,t_n\rangle$, then by (935.24), we know that $\forall x_1\ldots x_n\neg Fx_1\ldots x_n$ and $\forall x_1\ldots x_n\neg Gx_1\ldots x_n$ are modally strict theorems. Hence $\Box\forall x_1\ldots x_n(Fx_1\ldots x_n \equiv Gx_1\ldots x_n)$. *A fortiori*, $F = G \rightarrow \Box\forall x_1\ldots x_n(Fx_1\ldots x_n \equiv Gx_1\ldots x_n)$. So, by the contrapositive and modal negation, $\Diamond\neg\forall x_1\ldots x_n(Fx_1\ldots x_n \equiv Gx_1\ldots x_n) \rightarrow F \neq G$. Such facts prove useful in the proofs of the theorems in (.6).

The remainder of the definitions, theorems, and remarks in Section 9.10.2, i.e., items (196) – (226), will not be reviewed here. Most of these will not play a role in the applications of typed object theory. If any of these type-theoretic versions are needed in what follows, we'll state and prove them (possibly in a footnote) at the appropriate time.

Nevertheless, it may be of interest to determine which theorems of second-order object theory hold in complete generality in typed object theory. One will find cases where the proof of the second-order version appeals to a theorem that doesn't transfer to type theory in complete generality. For example, the proof of (199.7), i.e., $\overline{p} = \neg p$, appeals to (111.1), i.e., $[\lambda\,\varphi] = \varphi$. But the latter doesn't hold unconditionally in typed object theory. Instead, we have theorem (950.2), i.e., that $[\lambda\,\varphi] = \varphi$ provided φ is a non-basic formula. But, as it turns out, this is sufficient to prove $\overline{p} = \neg p$.[399] The reader should therefore check to see, when reformulating second-order theorems in type theory, whether alternative proofs are available in those cases where the second-order proofs rely on theorems that aren't completely general.

15.4.6 The Typed Theory of Objects

In typed object theory, We start with the theory of abstract objects, since we shall need the fact that there are distinct, but indiscernible higher-order abstract objects of each type t when we discuss the ordinary objects of type t.

[399] To see the proof, assume that definition of relation negation (196) has been restated type-theoretically. Then we know $\overline{p} = [\lambda\,\neg p]$. Moreover, by (950.2), we know $[\lambda\,\neg p] = \neg p$, since $\neg p$ is non-basic formula. So by the transitivity of identity (117.3), $\overline{p} = \neg p$.

Abstract Objects of Every Type

(**962**) **Theorems:** Some Identity and Existence Principles Governing Abstracta. The following theorems are typed versions of the theorems in (245) – (247). But (.1) and (.2) below also generalize theorems (961.3) and (961.4) to any type. Let t be any type, x and y be variables of type t, $A!$ be our defined constant of type $\langle t\rangle$, and F be a variable of type $\langle t\rangle$. Then we have (.1) if x and y encode the same properties, then necessarily they encode the same properties; and (.2) abstract objects x and y that encode the same properties are identical:

(.1) $\forall F(xF \equiv yF) \to \Box\forall F(xF \equiv yF)$

(.2) $(A!x \,\&\, A!y) \to (\forall F(xF \equiv yF) \to x = y)$

It is also provable that (.3) whenever an object x of any type encodes a property that object y of the same type fails to encode, or vice versa, x and y are distinct:

(.3) $(\exists F(xF \,\&\, \neg yF) \vee \exists F(yF \,\&\, \neg xF)) \to x \neq y)$

From axiom (935.31) it follows that (.3) if x encodes a property, then x is abstract:

(.4) $\exists F xF \to A!x$

The converse fails because there exists an abstract *null* object that encodes no properties. We introduce this object below.

Moreover, (.5) every property (ordinary or abstract) is encoded by some object; and, where G is a variable of type $\langle t_1,\ldots,t_n\rangle$, for any types $t_1,\ldots,t_n$, (.6) every relation (ordinary or abstract) is encoded By some objects.

(.5) $\forall F\exists x\, xF$

(.6) $\forall G\exists x_1 \ldots \exists x_n(x_1 \ldots x_n G)$ $\qquad (n \geq 2)$

In addition, given any relation term Π of type $t \neq \langle\,\rangle$, it follows that Π is encoded by some objects just in case there is a relation that is identical to Π:

(.7) $\exists x\, x\Pi \equiv \exists F(F = \Pi)$, where Π is any unary relation term of type $\langle t\rangle$ in which x and H don't occur free

(.8) $\exists x_1 \ldots \exists x_n(x_1 \ldots x_n\Pi) \equiv \exists G(G = \Pi)$, where Π is any n-ary relation term (a) having type $\langle t_1,\ldots,t_n\rangle$ $(n \geq 2)$, and (b) in which $x_1,\ldots,x_n$ and G don't occur free

We may refer to the above by type-theoretic numbers alone; we need not additionally cite their corresponding theorems in (245) – (247).

(**963**) **Theorems and Metadefinition:** Strengthened Comprehension for Abstracta and Significant Canonical Descriptions. Where t is any type, x is a variable of type t, $A!$ has type $\langle t \rangle$, F a variable having type $\langle t \rangle$, and φ is any formula in which x doesn't occur free, then the comprehension and identity principles for abstract object imply that (.1) there exists a unique abstract object x that encodes exactly the properties F such φ:

(.1) $\exists! x(A!x \,\&\, \forall F(xF \equiv \varphi))$

Now where t, x, $A!$, F, and φ are as described above, we say:

(.2) A description is *canonical* if and only if it has form $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$

It then follows that canonical descriptions are significant:

(.3) $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))\downarrow$

This holds for canonical descriptions of every type, i.e., canonical descriptions in which x has any type t. Now where y is also a variable of type t, then by reasoning analogous to the proof of theorem (255), it also follows that (.4) if y is the abstract object that encodes just the properties such that φ, y exemplifies being abstract:

(.4) $y = \imath x(A!x \,\&\, \forall F(xF \equiv \varphi)) \rightarrow A!y$

Next, we note that:

(.5) The definitions, discussion, and theorems governing canonical and *strictly* canonical abstract objects in (256) – (262) all carry over into typed object theory, appropriately typed. If and when these are needed, we:

- refer to the typed versions of the theorems in (256) as (963.5) [256.1]★ and (963.5) [256.2]★, respectively,
- refer to typed version of theorem (258) as (963.5) [258],
- refer to the typed versions of the theorems in (259) as (963.5) [259.1] and as (963.5) [259.2], respectively,
- refer to the definition of *rigid condition* on the variable α (260.1) and *strictly canonical description* (260.2) as (963.5) [260.1] and as (963.5) [260.2], respectively, and
- refer to the typed versions of the theorems in (261) as (963.5) [261.1] – (963.5) [261.3], respectively.

Exercise: The Null and Universal Objects. Define the null and universal objects of type t by typing the definitions in (263). Then (a) prove typed versions of the theorems in (264), (b) define *the null object* $\boldsymbol{a}^t_\varnothing$ for type t, and universal object $\boldsymbol{a}^t_V$ for type t, and (c) prove typed version of the theorems in (266).

(**964**) **Theorems:** Descriptions of Type $\langle\rangle$. By (963.3) and (940.1), it follows that every formula other than a non-canonical description of type $\langle\,\rangle$ is significant:

(.1) $\varphi\downarrow$, for any formula φ other than non-canonical description of type $\langle\,\rangle$

It also follows, where p is any variable of type $\langle\,\rangle$, that (.2) the *existence* of the proposition that φ does not imply the *truth* of the proposition that φ:

(.2) $\neg(\imath p\varphi\downarrow \rightarrow \imath p\varphi)$

(**965**) **Theorems:** Distinct Higher-order Abstracta that are Indistinguishable and Higher-Order Kirchner Theorems. The theorems about the granularity of relations and non-Leibnizian character of abstract objects hold in type theory:

(.1) Typed versions of the theorems in (268) are derivable. We refer to these as (965.1) [268.1] – (965.1) [268.3].

A typed version of (269) is also derivable. Where t be any type, x and y be variables of type t, $A!$ have type $\langle t\rangle$, and F be a variable of type $\langle t\rangle$, we have:

(.2) $\exists x\exists y(A!x \;\&\; A!y \;\&\; x\neq y \;\&\; \forall F(Fx \equiv Fy))$

Moreover, the typed version of the Kirchner Theorem (271) and its Corollary (272) are theorems. We recast only the general version (271.2) and its Corollary (272.2) in typed form. Where $t_1,\ldots,t_n$ are any types, $x_1,\ldots,x_n$ and $y_1,\ldots,y_n$ are, respectively, variables of types $t_1,\ldots,t_n$, respectively, and F is a variable of type $\langle t_1,\ldots t_n\rangle$, we have:

(.3) $[\lambda x_1\ldots x_n\,\varphi]\downarrow \equiv$
$\Box\forall x_1\ldots\forall x_n\forall y_1\ldots\forall y_n(\forall F(Fx_1\ldots x_n \equiv Fy_1\ldots y_n) \rightarrow (\varphi \equiv \varphi^{y_1,\ldots,y_n}_{x_1,\ldots,x_n}))$,
provided $y_1,\ldots,y_n$ don't occur free in φ. $(n\geq 1)$

(.4) $[\lambda x_1\ldots x_n\,\varphi]\downarrow \rightarrow$
$\forall x_1\ldots\forall x_n\forall y_1\ldots\forall y_n(\forall F(Fx_1\ldots x_n \equiv Fy_1\ldots y_n) \rightarrow \Box(\varphi \equiv \varphi^{y_1,\ldots,y_n}_{x_1,\ldots,x_n}))$,
provided $y_1,\ldots,y_n$ don't occur free in φ $(n\geq 1)$

And the consequences of the Corollary to the Kirchner Theorem, (272.3) – (272.5) hold in typed form. We prove here only the typed version of the general corollary (272.5). Where x,y are variables of type t, z is a variable of type t', and F is a variable of type $\langle t\rangle$, we have:

(.5) $[\lambda z\,\varphi]\downarrow \rightarrow (\forall F(Fx \equiv Fy) \rightarrow [\lambda z\,\varphi] = [\lambda z\,\varphi^y_x])$, provided none of the free occurrences of x in φ are in encoding position.

The proof of (.5) is given in the Appendix; as far as I can tell, no special measures are needed when adapting proofs of the second-order versions of (.1) – (.4) to typed object theory.

Finally, we prove a new theorem that arises only in typed object theory. Where t is any type, x is a variable of type t and F is a variable of type $\langle t\rangle$, that:

(.6) $\neg\forall x([\lambda F\, xF]\downarrow)$

Since F in xF occurs in encoding position (9.1), the λ in $[\lambda F\, xF]$ binds a variable in encoding position in the matrix. This theorem explains why (935.5) doesn't assert $[\lambda F\, xF]\downarrow$.

Ordinary Objects of Every Type

We first work our way through some theorems that will justify our introduction of *identity with respect to ordinary objects* as a defined relation, by cases, for every type t. As we shall see, one of the key cases of the definition of $=_E$ in (967.3). In that definition, the definiens is not a core λ-expression and so isn't guaranteed to denote by axiom (935.5). But we establish that it has a denotation by other means. We do this in (967.2), which requires the preliminary lemmas proved in (966).

(**966**) **Theorems:** Facts About Ordinary Properties. Where t is any type, x is a variable of type t, F is a variable of type $\langle t\rangle$, and $O!$ is a constant of type $\langle\langle t\rangle\rangle$, it follows that (.1) *being an ordinary property that x encodes* exists:

(.1) $[\lambda F\, O!F \,\&\, xF]\downarrow$

If we also use G as a variable of type $\langle t\rangle$ and $\mathcal{H}$ as a variable of type $\langle\langle t\rangle\rangle$, then it follows that (.2) ordinary properties F and G are necessarily encoded by the same objects if and only if they necessarily exemplify the same properties of properties, and that (.3) ordinary properties that necessarily exemplify the same properties of properties are identical:

(.2) $(O!F \,\&\, O!G) \rightarrow (\Box\forall x(xF \equiv xG) \equiv \Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G))$

(.3) $(O!F \,\&\, O!G) \rightarrow (\Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G) \rightarrow F = G)$

(.1) proves useful when we turn to the theory of identity with respect to ordinary objects, whereas (.2) offers an insight about the identity of ordinary properties, as defined in (933.10) – it makes no difference whether we define $F = G$ for ordinary properties in terms of being necessarily encoded by the same objects or in terms of necessarily exemplifying the same properties of properties. (.3) tells us that higher-order ordinary properties are obey the Leibniz law of the identity of indiscernibles.

(**967**) **Definitions and Theorems:** The Relation of Identity$_E$ on the Ordinary Objects of Any Type. Our goal is now to define identity for ordinary objects of type t, i.e., define $=_E$ as relation of type $\langle t, t\rangle$, for any type t. We do this by cases, where the four cases of the definition occur in (967.1), (967.3), (967.5), and (967.7).

It is easy to define $=_E$ as a relation among individuals, i.e., having type $\langle i,i\rangle$. Where x and y are variables of type i, $O!$ has type $\langle i\rangle$, and F is a variable of type $\langle i\rangle$:

(.1) $=_E =_{df} [\lambda xy\, O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)]$

Clearly, the definiens is a core λ-expression and hence significant. So we know that the relation $=_E$ among individuals exists.

Moreover, we can now also define $=_E$ as a relation for every other type $t \neq i$. First, we define $=_E$ as a relation with respect to any higher-order properties of the same type, i.e., as a relation with type $\langle\langle t\rangle,\langle t\rangle\rangle$, where t is any type. To do this, let t be any type, let x range over objects of type t, let F and G be variables ranging over properties having type $\langle t\rangle$, and let $O!$ have type $\langle\langle t\rangle\rangle$. Then it follows from theorem (966.2) and axiom (935.28) that (.2) *being ordinary properties F and G that are necessarily encoded by the same objects* exists:

(.2) $[\lambda FG\, O!F \,\&\, O!G \,\&\, \Box\forall x(xF \equiv xG)]\downarrow$

Given (.2), we may define identity$_E$ as a relation between higher-order properties. Where t is any type, and F and G are variables of type $\langle t\rangle$, we define:

(.3) $=_E =_{df} [\lambda FG\, O!F \,\&\, O!G \,\&\, \Box\forall x(xF \equiv xG)]$

In this definition, $=_E$ is a relation having type $\langle\langle t\rangle,\langle t\rangle\rangle$, for any type t.

We can now use the relation $=_E$ among entities with a property type of the form $\langle t\rangle$ to define $=_E$ as a relation among entities having an n-ary relation type $\langle t_1,\ldots,t_n\rangle$, for $n \geq 2$. That is, we now define $=_E$ as a relation of type $\langle\langle t_1,\ldots,t_n\rangle,\langle t_1,\ldots,t_n\rangle\rangle$ ($n \geq 2$), and we do this in terms of the various relations $=_E$ having types of the form $\langle\langle t\rangle,\langle t\rangle\rangle$, as these were defined in (.3). Let:

- $x_1,\ldots,x_n$ be variables of type $t_1,\ldots,t_n$, respectively,
- F and G be variables of type $\langle t_1,\ldots,t_n\rangle$, and
- $O!$ be a constant of type $\langle\langle t_1,\ldots,t_n\rangle\rangle$

Then where the first occurrrence of $=_E$ has type $\langle\langle t_1\rangle,\langle t_1\rangle\rangle$, the second occurrrence of $=_E$ has type $\langle\langle t_2\rangle,\langle t_2\rangle\rangle$, and …, and the nth occurrence of $=_E$ has type $\langle\langle t_n\rangle,\langle t_n\rangle\rangle$, the following claim is an instance of axiom (935.5) that asserts the significance of a core λ-expression:

(.4) $[\lambda FG\, O!F \,\&\, O!G \,\&\, \forall x_2\ldots\forall x_n([\lambda x_1\, Fx_1\ldots x_n] =_E [\lambda x_1\, Gx_1\ldots x_n]) \,\&$
$\forall x_1\forall x_3\ldots\forall x_n([\lambda x_2\, Fx_1\ldots x_n] =_E [\lambda x_2\, Gx_1\ldots x_n]) \,\&\ldots\&$
$\forall x_1\ldots\forall x_{n-1}([\lambda x_n\, Fx_1\ldots x_n] =_E [\lambda x_n\, Gx_1\ldots x_n])]\downarrow$

Thus, using the same typing scheme, we may define $=_E$ with respect to relations as a binary relation with type $\langle\langle t_1,\ldots,t_n\rangle,\langle t_1,\ldots,t_n\rangle\rangle$, for any types $t_1,\ldots,t_n$ ($n \geq 2$), as follows:

(.5) $=_E =_{df} [\lambda FG\, O!F \,\&\, O!G \,\&\, \forall x_2 \ldots \forall x_n([\lambda x_1\, Fx_1 \ldots x_n] =_E [\lambda x_1\, Gx_1 \ldots x_n]) \,\&$
$\forall x_1 \forall x_3 \ldots \forall x_n([\lambda x_2\, Fx_1 \ldots x_n] =_E [\lambda x_2\, Gx_1 \ldots x_n]) \,\&\ldots\&$
$\forall x_1 \ldots \forall x_{n-1}([\lambda x_n\, Fx_1 \ldots x_n] =_E [\lambda x_n\, Gx_1 \ldots x_n])])$

Finally, we want to define $=_E$ with respect to propositions, i.e., as a relation of type $\langle\langle\,\rangle,\langle\,\rangle\rangle$. Let x be a variable of type i, p and q be variables of type $\langle\,\rangle$, and let $O!$ have type $\langle\langle\rangle\rangle$. Then the following is an instance of axiom (935.5), where $=_E$ is a relation with type $\langle\langle i\rangle,\langle i\rangle\rangle$:

(.6) $[\lambda pq\, O!p \,\&\, O!q \,\&\, [\lambda x\, p] =_E [\lambda x\, q]]\!\downarrow$

So we may define $=_E$ as a relation with type $\langle\langle\,\rangle,\langle\,\rangle\rangle$, i.e., as a relation among propositions, as follows:

(.7) $=_E =_{df} [\lambda pq\, O!p \,\&\, O!q \,\&\, [\lambda x\, p] =_E [\lambda x\, q]]$

Thus, we have defined $=_E$ as a relation of type $\langle t,t\rangle$, for every type t, since it has been defined when t is the type i, when t is a type of the form $\langle t'\rangle$, when t is a type of the form $\langle t'_1,\ldots,t'_n\rangle$, and when t is the type $\langle\,\rangle$. Thus, we have:

(.8) $=_E\!\downarrow$, where $=_E$ has type $\langle t,t\rangle$, for any type t

(**968**) **Theorems:** Identity$_E$ is an Equivalence Relation on Ordinary Objects of Type t. Where x,y,z are objects of any type t, it now follows that $=_E$ is (.1) reflexive with respect to the ordinary objects of type t, (.2), symmetric, and (.3) transitive:

(.1) $O!x \to x =_E x$

(.2) $x =_E y \to y =_E x$

(.3) $(x =_E y \,\&\, y =_E z) \to x =_E z$

15.4.7 Discernible Objects Of Every Type

(**969**) It is straightforward to type all the definitions and theorems in (273) of Section 9.11.3. The λ-expression that is asserted to be significant in (273.1):

$$[\lambda x\, \Box\forall y(y \neq x \to \exists F \neg(Fy \equiv Fx))]$$

is provably significant whenever we let x and y be of any type t and F be of type $\langle t\rangle$. So the unary relation term $D!$ introduced in in (273.2) as a term of type $\langle t\rangle$, for any type t. Typed versions of theorems (273.3) – (273.12) are easily established by analogous reasoning. Principles (273.13) – (273.16), which establish (or help establish) the significant of:

$$[\lambda x\, D!x \,\&\, \varphi]$$

$[\lambda x_1 \ldots x_n\, D!x_1 \,\&\, \ldots \,\&\, D!x_n \,\&\, \varphi]$

$[\lambda xy\, D!x \,\&\, D!y \,\&\, x = y]$

all hold when the variables bound by the λ have arbitrary types. For example, when $x_1, \ldots, x_n$ have types $t_1, \ldots, t_n$, and the first occurrence of $D!$ has type $\langle t_1 \rangle$, ..., and the n-th occurrence of $D!$ has type $\langle t_n \rangle$, the λ-expression $[\lambda x_1 \ldots x_n\, D!x_1 \,\&\, \ldots \,\&\, D!x_n \,\&\, \varphi]$ is provably significant. Of course, x and y must have the same type if $[\lambda xy\, D!x \,\&\, D!y \,\&\, x = y]$ is to be significant, or even well-formed.

Moreover, the binary relation $=_D$ can be defined as a relation having type $\langle t, t \rangle$, for any type t, since its definiens $[\lambda xy\, D!x \,\&\, D!y \,\&\, x = y]$, is significant when x and y both are of any type t. This means that the typed negation of this relation, i.e., $\overline{=_D}$, which we may write as $\neq_D$, is similarly well-defined as a relation of type $\langle t, t, \rangle$, for any type t. Thus, principles (273.18) – (273.33) all hold in their typed versions, as do the final two theorems (273.34) and (273.35).

In what follows, we therefore reference the typed versions of these principles as (969) [273.1] – (969) [273.35].

15.4.8 Propositional Properties

(**970**) **Remark:** Propositional Properties in Typed Object Theory. A *propositional* property was defined in second-order object theory as any property F such that for some state of affairs p, F is identical to the property $[\lambda x\, p]$ (275). After proving a number of theorems about these properties in Section 9.12, we then investigated a number of abstract individuals that encode only propositional properties. These include truth-values (286), situations (467), possible worlds (512), world-indexed truth-values (555), impossible worlds (577), moments of time (588), and stories (592). From the point of view of type theory, these abstract entities are *individuals* that encode propositions by way of encoding propositional properties. They are individuals because they are not *predicable* entities. Insofar as we wish to study these entities in typed object theory, we no longer need to limit ourselves to the propositions that are expressible in second-order object theory. Instead, we can consider propositions (i.e., states of affairs) that are expressible in typed object theory, including any state of affairs expressible in terms of properties and relations of higher type.

The net effect, however, is that the propositional properties built from propositions expressible in terms of properties and relations of higher type, can remain properties of type $\langle i \rangle$. That is, we may define, where x is a variable of type i (i.e., ranging over individuals) and F is a variable of type $\langle i \rangle$ (ranging over properties of individuals), the following definition takes on new significance:

$$Propositional(F) \equiv_{df} \exists p(F = [\lambda x\, p])$$

The new significance arises from the fact that the variable p now ranges over a domain of propositions that include propositions about entities of higher type than those found in second-order object theory. Nevertheless, the theorems governing propositional properties studied in Section 9.12 can all be transferred unproblematically to typed object theory. We henceforth reference those as (970) [275] – (970) [281].

So, in so far as we are interested in theorizing about non-predicable abstract individuals in a typed context, we need only the above subtheory, which takes propositional properties to have type $\langle i \rangle$. Of course, one could build propositional properties of higher type, but for the purposes of this monograph, we shall not do so. We leave it to others to determine whether there is any interest in defining higher-order abstract objects that encode propositional properties of a type higher than $\langle i \rangle$.

15.5 Some Higher-Order Abstracta

As noted previously, in second-order object theory:

- the variables $x, y, z, \ldots$ range over entities having type i,
- the variables $F, G, H, \ldots$, range over entities having types $\langle t_1, \ldots, t_n \rangle$ such that $n \geq 1$ and $\forall t_j (1 \leq j \leq n \rightarrow t_j = i)$, and
- the variables $p, q, r, \ldots$ range over entities having type $\langle\, \rangle$.

Though Chapters 9 – 14 included various definitions and theorems about relations, properties, and propositions (i.e., about objects having types $\langle i, \ldots, i \rangle$, $\langle i \rangle$, and $\langle\, \rangle$, respectively), these chapters were primarily devoted to the study of abstract *individuals*, i.e., abstract objects having type i.[400] Intuitively, an entity is an individual if and only if it is not a predicable entity (i.e., if and only if it is not a relation, property, or proposition). So all of the abstract objects studied in Chapters 9 – 14 are abstract individuals.

The type-theoretic setting, however, gives rise to two questions: (1) How do the principles (definitions and theorems) governing the abstract individuals in second-order object theory change in the context of type theory? (2) Which of these principles for abstract individuals should be generalized to become principles for higher-order abstracta?

[400]These included: truth-values, extensions of propositions, extensions of properties (= natural classes = logical sets), directions, shapes, abstractions over equivalence conditions on properties and propositions, abstractions over equivalence relations on individuals, Forms, situations, possible worlds, world-indexed truth-values and extensions, impossible worlds, moments of time, fictional individuals, concepts, and natural numbers.

Question (1) is relatively straightforward. The definitions of the abstract individuals described in Chapters 9 – 14 effectively remain the same though, as one might expect (and as we shall see), the definitions take on greater significance. Moreover, the theorems governing these individuals can be preserved, though some theorem schemas must be limited to those formulas φ such that $\varphi\downarrow$, given that in typed object theory, there are formulas φ such that $\neg\varphi\downarrow$. But that said, the individuals we studied in these previous chapters remain, for the most part, governed by the principles discussed their respective chapters.

Question (2) is a more complex philosophical question. That's because in many cases, the entities we investigated are, given their nature, best conceived as individuals (i.e., non-predicable entities, i.e., having type i), while in other cases, the entities we investigated have natural higher-order versions when object theory is typed. That is, it is natural to generalize and investigate the higher-order versions of some, but all, of the objects studied in Chapters 9 – 14. So the discussion in what follows will be divided into two basic parts.

First, we discuss those objects which, given their nature, are (best conceived as) non-predicable individuals. These include truth-values, situations, possible worlds, world-relativized truth-values, impossible worlds, times, stories, fictional individuals, concepts, and natural numbers. As far as I can tell, there is no significant philosophical data that would motive one to investigate higher-order, predicable versions of these objects. Only Question (1) applies to them, and the considerations raised above in answer to this question will be discussed only in a general way in Section 15.5.1 below.

Second, we discuss those objects which have natural higher-order versions. For example, typed object theory now gives us the means to:

- identify the extensions of higher-order properties of type $\langle\langle t\rangle\rangle$ as abstract properties of type $\langle\langle t\rangle\rangle$, for any type t,

- identify higher-order abstractions over equivalence conditions and equivalence relations on higher-order properties of type $\langle\langle t\rangle\rangle$,

- identify, for each higher-order property F having type $\langle\langle t\rangle\rangle$, the Form of F as an abstract property having type $\langle t\rangle$, and

- identify fictional properties and relations of type $\langle i,\ldots,i\rangle$ as abstract properties and abstract relations of the type in question.

We'll discuss these objects in Sections 15.5.2 – 15.5.5, respectively. Along the way, in (972), we'll discuss the existence of world-indexed relations of every higher type

15.5.1 Abstractions Naturally Conceived as Individuals

In this section, we examine how some of the definitions and theorems governing a prototypical abstract individual fare in the type-theoretic setting. Our case study will focus on situations.

(**971**) **Definitions, Theorems, and Remark:** About Situations in Type Theory. We begin by typing definition (295) in the expected way, so that we may talk about an individual x *encoding* a proposition p. Where x is a variable of type i and p is a variable of type $\langle\,\rangle$, we say that x encodes p ('$x\Sigma p$') just in case x encodes *being such that* p:

(.1) $x\Sigma p \equiv_{df} x[\lambda y\, p]$

Then Where x is still a variable of type i and F a variable of type $\langle i\rangle$, the definition of a *situation* looks no different than the one given in (467.1):

(.2) $Situation(x) \ \equiv_{df} \ A!x \,\&\, \forall F(xF \rightarrow Propositional(F))$

And where x is a variable of type i, p is a variable of type $\langle\,\rangle$, and $x\Sigma p$ is defined as in (.1), the notion of *truth in* a situation has a natural type-theoretic counterpart:

(.3) $x \models p \ \equiv_{df} \ Situation(x) \,\&\, x\Sigma p$

Moreover, where s is a variable ranging over situations (and so a variable of type i), the *simplified* definition of a truth-value described in (486) can be formulated as:[401]

(.4) $TruthValueOf(s,p) \ \equiv_{df} \forall q(s \models q \equiv (q \equiv p))$

These definitions take on a new significance in typed object theory, because the variables p and q now range over a wider domain of propositions, namely, propositions constructed from the relations and properties having types higher than those in second-order object theory, i.e., having types $\langle t_1,\ldots,t_n\rangle$ such that $n \geq 1$ and $\exists t_j(t_1 \leq t_j \leq t_n \,\&\, t_j \neq i)$. This implies the existence of situations about higher-order objects. So for example, if s is a situation that encodes all and only the truths (i.e., $\forall p(s \models p \equiv p)$), then where q is a proposition not expressible in second-order object theory, e.g., the proposition that the relation $=_E$ (of type $\langle i,i\rangle$) exemplifies *being reflexive on the ordinary individuals* (i.e., $q = [\lambda F\, \forall x(O!x \rightarrow Fxx)] =_E$), then s encodes q. One can't formulate such a situation in second-order object theory.

[401]The original definition a truth-value in (286) is still formulable, of course. Let x and y be variables of type i, p and q be variables of type $\langle\,\rangle$, and let $A!$ and F have type $\langle i\rangle$. Then the following is well-formed:

$$TruthValueOf(x,p) \ \equiv_{df} \ A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q]))$$

But this is not as elegant as the simplified definition given in the text.

Thus, the unrestricted comprehension schema for situations expressed in (486.1) has a greater significance in typed object theory, for the quantifier $\forall p$ in the following principle ranges over propositions about objects of higher types:

(.5) $\exists s \forall p(s \models p \equiv \varphi)$, provided s doesn't occur free in φ

This is a theorem schema of typed object theory. Moreover, the theorem governing situation identity (474) can be typed as follows, in which the type of the variables are now the familiar ones:

(.6) $s = s' \equiv \forall p(s \models p \equiv s' \models p)$

(.2), (.3), (.5), and (.6), and the theorems they imply, offer a precise theory of situations expressible in higher-order metaphysics. These, then, are just a few examples of how to import the definitions and theorems governing the abstract individuals of second-order object theory that are best conceived as individuals in typed object theory.

With only one *proviso*, one may follow the examples (.1) – (.6) to complete the move to typed object theory. The *proviso* is that the theorems and theorem schemata transfer to typed object theory as long as they don't depend on theorem (104.2). This theorem asserts that $\varphi\downarrow$, for any formula φ. As we've seen, this claim doesn't transfer to typed object theory without restriction; the type-theoretic version is (940.1), which asserts that $\varphi\downarrow$ only when φ isn't a definite description of type $\langle\,\rangle$. As noted in the discussion of (931.1) and (950.1) – (950.3), the type-theoretic version of a second-order principle that depends on (104.2) isn't automatically disqualified as a theorem. But, the second-order theorem schemata that (a) govern the individuals we're now discussing, and (b) depend on (104.2), will often have type-theoretic versions that are conditionalized on formulas φ that are significant.[402]

To make the discussion more explicit, let's consider some examples of of second-order principles governing situations and possible worlds that require such a proviso. Theorem (511.3) has to be restricted in typed object theory. This theorem asserts $\forall p(s \models p \equiv p) \rightarrow ((s \models \forall \alpha\varphi) \equiv \forall \alpha(s \models \varphi))$, i.e., if a situation s encodes all and only the truths, then the universal claim every α is such that φ is true in s if and only if, for every α, φ is true in s. The proof depends on the fact that $\varphi\downarrow$ and $(\forall\alpha\varphi)\downarrow$, for every φ.[403] So, unless some other proof can be

[402] Thus, in cases where one can prove that a formula φ of the form $\imath p\psi$ has a denotation, it can be instantiated into the theorem schema in question and the antecedent of the conditional can be discharged. See below for examples.

[403] To see why the proof won't transfer to typed object theory, assume the antecedent, i.e.,

(ϑ) $\forall p(s \models p \equiv p)$

Now for the ($\rightarrow$) direction, assume $s \models \forall\alpha\varphi$. By GEN, it suffices to show $s \models \varphi$. Now by (940.1), we know $(\forall\alpha\varphi)\downarrow$. So we can instantiate (ϑ) to obtain $(s \models \forall\alpha\varphi) \equiv \forall\alpha\varphi$. Hence $\forall\alpha\varphi$, from which it follows that φ. But we can't now infer from this and (ϑ) that $s \models \varphi$ because φ might be a non-found, one cannot import (511.3) into typed object theory. It can be imported, however, if one conditionalizes the theorem on formulas φ such that $\varphi\downarrow$. Thus, in typed object theory, (511.3) becomes:

(.7) $\varphi\downarrow \rightarrow \big(\forall p(s \models p \equiv p) \rightarrow ((s \models \forall\alpha\varphi) \equiv \forall\alpha(s \models \varphi))\big)$

By contrast, (511.2) can be typed as is, despite the fact that its proof also appeals to (104.2). The formula for which (104.2) was used as justification, namely $(\neg(q \rightarrow r))\downarrow$, is provably a theorem of typed object theory, by (940.1).

Moreover, second-order theorems that depend, in turn, on (511.3), have to be checked carefully when we transfer them to type theory. Consider, for example, some theorems of world theory that depend on (511.3). The type-theoretic definition of a *possible world* recapitulates (512), though with somewhat greater significance given the expanded range of the quantifier $\forall p$:

(.8) $PossibleWorld(x) \equiv_{df} Situation(x) \,\&\, \Diamond\forall p(x \models p \equiv p)$

Now most of the main theorems governing possible worlds (543.1) and (543.2) transfer to type theory as theorems without restriction but with expanded significance. For example, where w is a variable of type i ranging over possible worlds defined type-theoretically, here are three such principles that will prove useful in what follows:

(.9) $w \models \neg p \equiv \neg w \models p$

(.10) $\Diamond p \equiv \exists w(w \models p)$

(.11) $\Box p \equiv \forall w(w \models p)$

Note, however, that the second-order theorems (545.5) and (545.6) that govern the definition of *possible world* both rest on (511.3). They assert, respectively, that the universal claim $\forall\alpha\varphi$ is true at a possible world w iff for every α, φ is true at w, and that the existential claim $\exists\alpha\varphi$ is true at a possible world w iff for some α, φ is true at w. We have to conditionalize these claims when we formulate their type-theoretic counterparts as follows, where the variable w again has type i and the variable x is a variable of any type:

(.12) $\varphi\downarrow \rightarrow \big((w \models \forall x\varphi) \equiv \forall x(w \models \varphi)\big)$

---

denoting description of type $\langle\,\rangle$. That is, we can't instantiate (ϑ) to obtain $(s \models \varphi) \equiv \varphi$ and, without that, we can't infer $s \models \varphi$ from φ.

The $(\leftarrow)$ direction similarly fails. Assume $\forall\alpha(s \models \varphi)$. Then $s \models \varphi$. Now if we could infer $(s \models \varphi) \equiv \varphi$ from (ϑ), we could conclude φ, which by GEN yields $\forall\alpha\varphi$ (α is not free in any assumption). Since $(\forall\alpha\varphi)\downarrow$ (see the previous paragraph), we would then be able to instantiate it for $\forall p$ in (ϑ); from the resulting biconditional, $(s \models \forall\alpha\varphi) \equiv \forall\alpha\varphi$, we could then conclude $s \models \forall\alpha\varphi$. But, alas, we may not invoke this chain of reasoning because we may not infer $(s \models \varphi) \equiv \varphi$ from (ϑ). The chain of reasoning is valid only if $\varphi\downarrow$, and understandably so.

(.13) $\varphi\downarrow \rightarrow \big((w \models \exists x\varphi) \equiv \exists x(w \models \varphi)\big)$

We leave it to the reader to find other examples where the principles governing the abstract individuals of second-order object theory, when imported into typed object theory, have to be constrained so as not to apply to non-denoting formulas.

(**972**) **Theorems:** The Existence of World-Indexed Relations of Higher Type. Though the present section has focused on abstracta naturally conceived as individuals, it is worth remarking on the fact that world-indexed relations of every higher type *provably* exist. This stands in contrast to Williamson 2013 (237), in which such relations are stipulated to exist in the semantics of his system of intensional modal logic. In the metalanguage of his higher-order metaphysics, Williamson (a) introduces an additional basic type, namely, a primitive type w for possible worlds, and (b) constructs higher-order types $\langle t_1, \ldots, t_n, w\rangle$ for world-indexed relations among objects of type $\langle t_1, \ldots, t_n\rangle$. But in higher-order object theory, both possible worlds and world-indexed relations of higher type are defined, and their existence derived, solely using the resources of the object language and the axioms of the system.

We establish the existence of world-indexed relations of every higher type in the following sequence of principles, in which $t_1, \ldots, t_n$ are any types; the variables $x_1, \ldots, x_n$ have types $t_1, \ldots, t_n$, respectively; w has type i; and F has type $\langle t_1, \ldots, t_n\rangle$. It follows that (.1) *being objects* $x_1, \ldots, x_n$ *such that, at* w, $x_1, \ldots, x_n$ *exemplify* F exists. Then we say (.2) the property *being* F *at* w ('F_w') is, by definition, the property *being objects* $x_1, \ldots, x_n$ *such that, at* w, $x_1, \ldots, x_n$ *exemplify* F. It then follows that (.3) for every relation F and possible world w, *being* F *at* w exists:

(.1) $[\lambda x_1 \ldots x_n\, w \models F^n x_1 \ldots x_n]\downarrow$ $\quad (n \geq 0)$

(.2) $F_w =_{df} [\lambda x_1 \ldots x_n\, w \models F^n x_1 \ldots x_n]$ $\quad (n \geq 0)$

(.3) $\forall F \forall w(F_w\downarrow)$ $\quad (n \geq 0)$

Clearly, F_w is a property of type $\langle t_1, \ldots, t_n\rangle$, and since the theorem holds for any types $t_1, \ldots, t_n$, we've established the existence of world-indexed relations of every higher type. Note that we can instantiate (.3) to any n-ary relation of type $\langle t_1, \ldots t_n\rangle$, including abstract ones. In the case where G is an abstract relation, G_w is an *ordinary* relation, given axiom (935.22). Moreover, by appealing to the typed facts that possible worlds are coherent and that $\Box p \equiv \forall w(w \models p)$, we may prove (.4) if G is abstract, then G_w is necessarily unexemplified, i.e., where $A!$ has type $\langle\langle t_1, \ldots, t_n\rangle\rangle$:

(.4) $A!G \rightarrow \Box\neg\exists x_1 \ldots \exists x_n G_w x_1 \ldots x_n$

We turn next to those second-order definitions of abstract individuals that can be naturally generalized to higher-order definitions of abstract objects of every type.

15.5.2 Extensions of Higher-Order Properties

The first example concerns the extensions of higher-order properties. In (312), we said that an individual x is an *extension of* a property G just in case x is abstract, G exists, and x encodes all and only the properties materially equivalent to G. This can be generalized as follows. For any type t, we may say that an extension of a higher-type property G having type $\langle t\rangle$ is an abstract object of type t that encodes just the properties having type $\langle t\rangle$ that are materially equivalent to G.

(**973**) **Definitions and Theorems:** The Extensions of (Higher-Order) Properties. Let t be any type; x, z be variables of type t; F, G be variables of type $\langle t\rangle$; and $A!$ be a defined constant of type $\langle t\rangle$. Then consider the result of importing definition (312) into typed object theory by typing it as follows:

$$(.1)\quad \left.\begin{array}{c} \mathit{ExtensionOf}(x,G) \\ \mathit{ClassOf}(x,G) \end{array}\right\} \equiv_{df} A!x \,\&\, G\!\downarrow \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz))$$

So, for example, suppose t is $\langle i\rangle$. If G is a higher-order property having type $\langle\langle i\rangle\rangle$, then an abstract property x having type $\langle i\rangle$ is an extension of G just in case x encodes exactly the type $\langle\langle i\rangle\rangle$ properties materially equivalent to G (w.r.t. exemplification).

But now consider the general case again, where t is arbitrary. Though we are considering here just the extensions of properties having type $\langle t\rangle$, we can now consider the extensions of both ordinary and abstract properties of this type. Consider this fact in connection with the following two cases:

- G is an exemplified property of type $\langle t\rangle$
- G is an unexemplified property of type $\langle t\rangle$

In the first case, we have: (.2) any object x that is an extension of an exemplified property G having type $\langle t\rangle$ encodes only ordinary properties having type $\langle t\rangle$. I.e., where $O!$ is our defined constant of type $\langle\langle t\rangle\rangle$:

$$(.2)\quad \big(\exists zGz \,\&\, \mathit{ExtensionOf}(x,G)\big) \to \forall F(xF \to O!F)$$

Clearly, if G is an exemplified property having type $\langle t\rangle$, and axiom (935.24) guarantees that abstract properties of every type are unexemplified, the only properties having type $\langle t\rangle$ that are materially equivalent to G are ordinary properties.

In the second case, we have: (.3) if x is an extension of an unexemplified property of type $\langle t\rangle$, then x of G encodes every abstract property of type $\langle t\rangle$; and (.4) if x is an extension of an unexemplified *ordinary* property of type $\langle t\rangle$ and y is an extension of any *abstract* property of that type, then x is identical to y. I.e., where H is another variable of type $\langle t\rangle$:

(.3) $\big(\neg\exists zGz \,\&\, ExtensionOf(x,G)\big) \rightarrow \forall F(A!F \rightarrow xF)$

(.4) $\big(O!G \,\&\, \neg\exists zGz \,\&\, ExtensionOf(x,G) \,\&\, A!H \,\&\, ExtensionOf(y,H)\big) \rightarrow x = y$

Moreover, since abstract properties are unexemplified (again, by axiom (935.24)), we have that p(.5) an extension x of an abstract property G of type $\langle t\rangle$ encodes every abstract property of type $\langle t\rangle$:

(.5) $\big(A!G \,\&\, ExtensionOf(x,G)\big) \rightarrow \forall F(A!F \rightarrow xF)$

(**974**) **Theorems and Definitions:** Unique Extensions for (Higher-Order) Properties. Let t be any type, x and y be variables of type t, and G be a variable of type $\langle t\rangle$. Then it follows that: (.1) there is a unique extension of G of type t; and (.2) the extension of G exists:

(.1) $\exists!xExtensionOf(x,G)$

(.2) $\imath xExtensionOf(x,G)\downarrow$

Moreover, we may define, (.3) ϵG is defined as the extension of G, and, where y is an object of type t, (.4) y is an element of x holds by definition whenever y exemplifies some property G of which x is an extension:

(.3) $\epsilon G =_{df} \imath xExtensionOf(x,G)$

(.4) $y \in x \equiv_{df} \exists G(ExtensionOf(x,G) \,\&\, Gy)$

15.5.3 Abstractions over Higher-Order Equivalences

(**975**) **Metadefinition, Definitions, and Theorems**: Abstraction on Higher-Order Equivalence Conditions and Equivalence Relations. Let $t_1,\ldots,t_n$ be any types and let φ be any formula in which there are free of occurrences of the *two* distinct relation variables having type $\langle t_1,\ldots,t_n\langle$ (for some n). Suppose we've distinguished these free variables from any other free variables that may occur in φ, and that we may refer to one of these distinguised variables as 'the first' if it has the first free occurrence in φ (and refer to the other as 'the second'). Then where α and β are *any* two n-ary relation variables of type $\langle t_1,\ldots,t_n\langle$, let us write $\varphi(\alpha,\beta)$ for the result of simultaneously substituting α for all the free occurrences of the first distinguished free variable in φ and substituting β for all the free occurrences of the second distinguished free variable in φ. Thus, if

φ happens to have α and β as the two distinguished free variables, then $\varphi(\alpha, \beta)$ just is φ. Given this notational convention, we say:

(.1) **Equivalence Condition:** A formula φ with two distinct n-place relation variables is an *equivalence condition* on n-ary relations whenever the following are all provable:

$$\varphi(\alpha, \alpha) \qquad \text{(Reflexivity)}$$

$$\varphi(\alpha, \beta) \to \varphi(\beta, \alpha) \qquad \text{(Symmetry)}$$

$$\varphi(\alpha, \beta) \to (\varphi(\beta, \gamma) \to \varphi(\alpha, \gamma)) \qquad \text{(Transitivity)}$$

Whenever φ is an equivalence condition on relations of any type, and $\varphi(\alpha, \beta)$ holds, we say that α and β are φ-equivalent.

Now since the definition of φ-*AbstractionOf*(x, p) in (385.1) remains the same when p is a variable of type $\langle\,\rangle$, we need only examine definition (385.2) and consider the case where t is any type other than $\langle\,\rangle$, x is a variable of type t, F, G are variables of type $\langle t \rangle$, and $A!$ is the defined constant of type $\langle t \rangle$. That is, where ψ is any equivalence condition on properties having type $\langle t \rangle$, we may type definition (385.2) so that it stipulates that x is the ψ-abstraction of G if and only if x is abstract, G exists, and x encodes just the properties F that are ψ-equivalent to G:

(.2) $\psi\text{-}AbstractionOf(x, G) \ \equiv_{df} \ A!x \,\&\, G\!\downarrow \,\&\, \forall F(xF \equiv \psi(F, G))$

Then following this typing scheme, we have the following sequence of theorems and definitions, with H a variable as the same type as G:

(.3) $\exists! x(\psi\text{-}AbstractionOf(x, G))$

(.4) $\iota x(\psi\text{-}AbstractionOf(x, G))\!\downarrow$

(.5) $(\psi\text{-}AbstractionOf(x, G) \,\&\, \psi\text{-}AbstractionOf(y, H)) \to (x = y \equiv \psi(G, H))$

(.6) $\widehat{G}_{\psi} \ =_{df} \ \iota x(\psi\text{-}AbstractionOf(x, G))$

We may also derive, as a non-modally strict theorem, a Fregean biconditional that corresponds to (389)★, namely, $\widehat{F}_{\psi} = \widehat{G}_{\psi} \ \equiv \ \psi(F, G)$.

We next turn to abstractions over higher-order equivalence relations. Let t be any type, and let F be a binary relation variable having type $\langle t, t \rangle$. Then we say that F is an *equivalence* relation on objects of type t if and only if F is reflexive, symmetric, and transitive, i.e., where x, y, and z are variables of type t:

(.7) $Equivalence_t(F) \ \equiv_{df}$
$\forall x Fxx \,\&\, \forall x \forall y(Fxy \to Fyx) \,\&\, \forall x \forall y \forall z(Fxy \,\&\, Fyz \to Fxz)$

Clearly, where t is any type, x and y are variables of type t and F is a variable of type $\langle t\rangle$, we can establish that $[\lambda xy\,\forall F(Fx \equiv Fy)$ exists and is an equivalence relation$_t$, and that any relation term that satisfies the definition has is significant. So $Equivalence_t(F)$ is a restriction condition on relations, we may again henceforth use $\widetilde{R}$ as a restricted variable ranging over equivalence relations having type $\langle t,t\rangle$.

Exercise: Show that the theorems and definitions in (413) – (419) are generalizable, i.e., that (413) – (419), when typed using the above typing scheme, remain principles of typed object theory.

15.5.4 Forms

Exercise: Let t be any type; $x,y,z,\ldots$ be objects of type t; $F,G,H,\ldots$ be properties of type $\langle t\rangle$; and $A!$ be the defined property of *being abstract* having type $\langle t\rangle$. Then where $F \Rightarrow G$ is defined in the usual way as $\Box\forall x(Fx \rightarrow Gx)$, consider the following, typed versions of definitions (421) and (444):

$$ThinFormOf(x,G) \equiv_{df} A!x \,\&\, G\!\downarrow \,\&\, \forall F(xF \equiv F\!=\!G)$$

$$FormOf(x,G) \equiv_{df} A!x \,\&\, G\!\downarrow \,\&\, \forall F(xF \equiv G\!\Rightarrow\! F)$$

Now show that the work in Chapter 11 generalizes to type theory, i.e., that the definitions and theorems of Chapter 11 remain principles that govern the thin and thick Forms for properties of any type, i.e., govern the thin and thick Forms for any object having a type of the form $\langle t\rangle$.

Exercise: Since all abstract properties are unexemplified and necessarily so, what generalization can one draw about the properties encoded by the thick Form x of an abstract property G?

(**976**) **Remark:** Thin Forms as (First-Order) Representatives of Higher-Order Extensions. To prepare an example that shows how a thin Form of an abstract property of individuals can represent a class of higher-order properties of properties of individuals, let G be a variable of type $\langle i\rangle$, and let $\mathcal{K}$ be a property of properties of individuals, i.e., having type $\langle\langle i\rangle\rangle$. Then we know, by an appropriate instance of (974.1):

$$\exists! \textsc{g}(ExtensionOf(\textsc{g},\mathcal{K}))$$

Suppose P is the unique such property, so that we know $ExtensionOf(\textsc{p},\mathcal{K})$. Then where H is a variable of type $\langle i\rangle$ and $\mathcal{F}$ is a variable of type $\langle\langle i\rangle\rangle$, and $A!$ has type $\langle\langle i\rangle\rangle$, it follows *a fortiori* from definition (973) and the existence of $\mathcal{K}$ that:

$$A!\textsc{p} \,\&\, \forall\mathcal{F}(\textsc{p}\mathcal{F} \equiv \forall \textsc{h}(\mathcal{F}\textsc{h} \equiv \mathcal{K}\textsc{h}))$$

And by definition (974.4) and the fact that $ExtensionOf(P, \mathcal{K})$, we know that $H \in P \equiv \mathcal{K}H$.

But now note that we can represent P by an abstract individual, namely, the Thin Form of P, i.e., $\boldsymbol{a}_P$, where this is defined via the typed version of (423). Furthermore, we can *extend* the notion of membership so as to define a sense in which the thin Form of G ($\boldsymbol{a}_G$) has, as members, higher-order properties of properties of individuals. That is, we may say that $\mathcal{F}$ is an element of $\boldsymbol{a}_G$ if and only if G is an extension of some higher-order property and encodes $\mathcal{F}$:

$$\mathcal{F} \in \boldsymbol{a}_G \equiv_{df} \exists\mathcal{K}(ExtensionOf(G, \mathcal{K}) \,\&\, G\mathcal{F})$$

This gives us a mechanism by which an abstract individual such as $\boldsymbol{a}_P$ may represent a class of higher-order properties such as P.

15.5.5 Fictional Properties and Fictional Relations

In Chapter 12, Section 12.6, we analyzed stories and fictional individuals as abstract individuals. We are now in a position to analyze fictional properties, like *being a unicorn*, *being a hobbit*, and fictional relations, like *absolute simultaneity*, etc. We'll focus primarily on fictional properties and so, for the most part, leave the generalization to fictional relations as an exercise (the exception being a fictional relation of identity$_{\underline{s}}$ for those stories $\underline{s}$ that explicitly or implicitly assume a relation of identity).

(**977**) **Remark:** The Data. In Chapter 12, Section 12.6, we divided the data into four categories:

(A) pre-theoretic *truths* about stories and characters that would become falsehoods if prefaced by a locution of the form 'In the story' or 'According to the story';

(B) true claims about the world that are taken to be true when authors use them in the context of a story;

(C) pre-theoretic truths that have the form 'In the story σ, ...' or 'According to the story σ, ...'; and

(D) pre-theoretic judgments about what logically follows from the truths in (A), (B), and (C).

Now that we have the expressive power to talk about the properties of properties and the properties of relations, we can expand these categories to include a greater variety of examples. Thus, we have the following new data to analyze. In (A):

(.1) *Being a hobbit* originates in the *The Hobbit* novels.

(.2) *Being a hobbit* is a fictional property.

(.3) *Acromantula* is a fictional species of arachnid.

In (B):

(.4) Humans are mortal.

(.5) Gold is a precious metal.

Note that (.4) remains true when prefaced by the story operator "In *The Lord of the Rings*", and (.5) remains true when prefaced by the story operator "In the *Harry Potter* novels".

In (C):

(.6) According to *The Lord of the Rings*, *hobbits* are a species whose individuals are of short stature.

(.7) According to *The Lord of the Rings*, *mithril* is a kind of metal pieces of which are lightweight and strong.

In (D):

(.8) In *The Hobbit*, Bilbo is a hobbit.
Being a hobbit originates in *The Hobbit*.
Therefore, *being a hobbit* is a fictional property.

(.9) Peter Jackson made a movie about *being a hobbit*.
Hobbits are a fictional species.
Fictional species aren't real.
Therefore, Peter Jackson made a movie about a species that isn't real.

In light of the vast science fiction and fantasy literature, these examples suffice to demonstrate that that there is a wide range of data about fictional properties.

(**978**) **Assumptions, Definitions, Theorems:** Typing the Principles for Analyzing Fiction. We approach the analysis of the data by first typing the assumptions we introduced for analyzing fiction. The authorship relation ('A') (591) is a relation between ordinary individuals and abstract individuals (stories) and so has type $\langle i,i\rangle$. Thus, where x and y are variables of type i, we may reassert the two minimal principles for A (591.1) and (591.2) in their type-theoretic guise:

(.1) $Ayx \rightarrow E!y$ (**Assumption**)

(.2) $\exists yAyx \rightarrow \Diamond\neg\exists yAyx$ (**Assumption**)

Moreover, where $\underline{s}$ is a restricted variable ranging over stories (i.e., individuals defined below), p is a variable of type $\langle\,\rangle$ ranging over propositions, $\underline{s} \models p$ is defined in the usual way as $s\Sigma p$, i.e., $\underline{s}[\lambda y\, p]$, and $x_1,\ldots,x_n$ are distinct variables having types $t_1,\ldots,t_n$, respectively, we may reassert the two minimal principles for relevant entailment ($\Rightarrow_R$) in (591.3) and (591.4) in their type-theoretic guise:

(.3) $(\underline{s} \models p_1 \;\&\; \ldots \;\&\; \underline{s} \models p_n \;\&\; (p_1,\ldots,p_n \Rightarrow_R q)) \rightarrow \underline{s} \models q$ (**Assumption**)

(.4) $[\lambda x_1 \ldots x_n\, \varphi]\!\downarrow\; \rightarrow ([\lambda x_1 \ldots x_n\, \varphi]x_1 \ldots x_n \Leftrightarrow_R \varphi)$ (**Assumption**)

To minimize the clash of variables (the variable t was introduced in a previous chapter to range over times but is used in this chapter to range over types), we omit the formal statement of the principle that governs *before* (591.5), asserting only its typed version as follows:

(.5) *p temporally precedes q* just in case there are times t and t' such that t precedes t', p occurs at t but at no earlier time, and q occurs at t' but at no earlier time. (**Definition**)

Given assumptions (.1) – (.5), we may type all of the principles (definitions, theorems, axioms★, axioms, etc.) in Section 12.6.2, namely, (592) – (593), and (596) – (597). This is mostly routine (see below) and so we shall henceforth refer to the type-theoretic versions of these principles, respectively, as (978) [592] – (978) [593], and (978) [596] – (978) [597]. Thus, we may use (978) [592] to refer to the typed definition of a story as any non-null situation that is authored.

However, the principles (598) – 604 require a liberalization. Since our goal is to develop a theory of fictional properties and relations, it is important to type these principles so as to allow for higher-type entities to be characters of stories. We do this as follows.

The definition of *x is a character of* $\underline{s}$, has to be typed so that the x can be of higher type. So, where x is a variable of any type t, and F is a variable of type $\langle t\rangle$, the typed version of definition (598) becomes:

(.6) $CharacterOf(x,\underline{s}) \equiv_{df} x\!\downarrow \;\&\; \exists F(\underline{s} \models Fx)$

This has greater significance than (598), since it allows relations to be characters of stories. For example, we might use this to assert that *being a hobbit*, where this is a property having type $\langle i\rangle$, is a character of Tolkien's *The Hobbit*. Note that for stories that have indiscernible characters, (.6) should be revised in favor of (.19), which appears at the end of this discussion.

Similarly, where x is an object of some type t, we have (.7) *x originates in* story $\underline{s}$ just in case x is a character of $\underline{s}$ that is abstract and that is not a character of any story authored before $\underline{s}$. To formalize this, let x be a variable of any type t and $A!$ be a defined constant of type $\langle t\rangle$, we have:

(.7) $OriginatesIn(x,\underline{s}) \equiv_{df} CharacterOf(x,\underline{s})\ \&\ A!x\ \&$
$\forall y\forall z\forall \underline{s}'((Az\underline{s}' < Ay\underline{s}) \rightarrow \neg CharacterOf(x,\underline{s}'))$

For example, we may use this to assert that*being a hobbit* originates in *The Hobbit*. Thus, (.8) x is an *original character of* $\underline{s}$ just in case x originates in $\underline{s}$:

(.8) $OriginalCharacterOf(x,\underline{s}) \equiv_{df} OriginatesIn(x,\underline{s})$

It is straightforward, then, to type axiom (600) for identifying original characters as follows. Where x and y are variables of type t, $A!$ is a defined constant of type $\langle t\rangle$, and F is a variable of type $\langle t\rangle$, we assert, as an axiom, the following type-theoretic incarnation of (600), namely (.9) if object x is an original character of $\underline{s}$, then x encodes exactly those properties that x exemplifies according to $\underline{s}$:

(.9) $OriginalCharacterOf(x,\underline{s}) \rightarrow x = \imath y(A!y\ \&\ \forall F(yF \equiv\ \underline{s} \models Fy))$ (**Axiom**)

If we let x be a variable of type $\langle i\langle$, and suppose *being a hobbit* has that type, and the story *The Hobbit* (h) has type i, then as an applied instance of (.9), we have: if *being a hobbit* (H) is an original character of *The Hobbit*, then H is the abstract property that encodes just the properties of properties that H exemplifies according to *The Hobbit*, i.e.,

- $OriginalCharacterOf(H,\underline{h}) \rightarrow H = \imath y(A!y\ \&\ \forall F(yF \equiv\ \underline{h} \models Fy))$

Thus, properties may be original characters of stories. Where x is a variable of any type t, we may assert the following typed counterpart of (601), namely, (.10) x is a fictional character just in case it is an original character of some story:

(.10) $Fictional(x) \equiv_{df} \exists \underline{s} OriginalCharacterOf(x,\underline{s})$

For example, this definition lets us infer, from the premise that Raskolnikov is an original character of *Crime and Punishment*, the conclusion that Raskolnikov is a fictional character.

Next, where x is an object of type t and G is a property of type $\langle t\rangle$, we say that (.11), x is a *fictional* G just in case x is an original character of a story according to which x exemplifies G:

(.11) $Fictional\text{-}G(x) \equiv_{df} \exists \underline{s}(OriginalCharacterOf(x,\underline{s})\ \&\ \underline{s} \models Gx)$

Consider an instance of this definition when x has type $\langle i\rangle$, and G has type $\langle\langle i\rangle\rangle$. For example, let x be the property *being a unicorn* (U) and let G be the higher-order property of properties *being a species* (S). Then as an instance of (.11), we have:

- $Fictional\text{-}S(U) \equiv_{df} \exists \underline{s}(OriginalCharacterOf(U,\underline{s})\ \&\ \underline{s} \models SU)$

> *Being a unicorn* is a fictional species if and only if there exists a story $\underline{s}$ such that *being a unicorn* is an original character of $\underline{s}$ and in the story, *being a unicorn* is a species.

In light of the foregoing definitions, some interesting theorems follow. First, note that (.12) fictional objects are abstract, and (.13) fictional Gs are abstract. We may represent these type-theoretically if we let x be a variable of type t, F be a variable of type $\langle t\rangle$, and $A!$ have type $\langle t\rangle$:

(.12) $\mathit{Fictional}(x) \rightarrow A!x$

(.13) $\mathit{Fictional}\text{-}F(x) \rightarrow A!x$

So when x has type $\langle i\rangle$ and we instantiate it to the property *being a hobbit* ($\textsc{h}$), and F in (.13) is a variable of type $\langle\langle i\rangle\rangle$ and we instantiate it to the property *being a species* (S), we may instantiate (.12) and (.13), respectively, to derive:

- $\mathit{Fictional}(\textsc{h}) \rightarrow A!\textsc{h}$, i.e.,

  If *being a hobbit* is (a) fictional (property), it is (an) abstract (property).

- $\mathit{Fictional}\text{-}S(\textsc{h}) \rightarrow A!\textsc{h}$

  If *being a hobbit* is a fictional species, it is abstract.

Moreover, we may type theorem (603.1) so as to derive, where x has any type t:

(.14) $A!x \rightarrow \neg\exists y(\Diamond E!y \,\&\, y = x)$

Now suppose that when some philosophers talk about possible objects, they are referring to possibly concrete objects, and when they talk about possible properties, they are referring to possibly concrete properties. Then from the foregoing we obtain type-theoretic counterparts of (603.2) and (603.3); it follows immediately that, for an type, (.15) a fictional object is not identical with any possible object, and (.16) a fictional F is not identical with any possible F:

(.15) $\mathit{Fictional}(x) \rightarrow \neg\exists y(\Diamond E!y \,\&\, y = x)$

(.16) $\mathit{Fictional}\text{-}F(x) \rightarrow \neg\exists y\big(\Diamond(E!y \,\&\, Fy) \,\&\, y = x\big)$

So, as higher order examples, suppose x is again a variable of type $\langle i\rangle$ and instantiate it to the property *being a hobbit* (H), and instantiate F to the higher-order property *being a species* ($\mathcal{S}$), which has type $\langle\langle i\rangle\rangle$. There where $E!$ has type $\langle\langle i\rangle\rangle$, we have, as applied instances of (.15) and (.16), the following as theorems:

- $\mathit{Fictional}(\textsc{h}) \rightarrow \neg\exists \textsc{g}(\Diamond E!\textsc{g} \,\&\, \textsc{g} = \textsc{h})$

If *being a hobbit* is fictional, then it is not identical with any possibly concrete property.

- $Fictional\text{-}\mathcal{S}(H) \rightarrow \neg\exists G\big(\Diamond(E!G \;\&\; \mathcal{S}G) \;\&\; G = H\big)$

  If *being a hobbit* is a fictional species, then it is not identical with any possibly concrete species.

Now let's suppose that when Kripke talks about 'possible species' in 1972 [1980], he means to refer to possibly concrete species. Then, if we instead substitute the property of *being a unicorn* for x in (.16), we can derive of a key claim in Kripke 1972 [1980] from his assumption that unicorns are a mythical species. He concludes from this that we can't identify unicorns with any possible species:

> As to the metaphysical thesis, the argument basically is the following. Just as tigers are an actual species, so the unicorns are a mythical species. . . . there is no actual species of unicorns, and regarding the several distinct hypothetical species, with different internal structures (some reptilic, some mammalian, some amphibious), which would have the external appearances postulated to hold of unicorns in the myth of the unicorn, one cannot say which of these distinct mythical species would have been the unicorns. If we suppose, as I do, that the unicorns of the myth were supposed to be a particular species, but that the myth provides insufficient information about their internal structure to determine a unique species, then there is no actual or possible species of which we can say that it would have been the species of unicorns. Kripke 1972 [1980, 156–7])

In the present theory, the claim that *being a unicorn* is not identical with any possible species, i.e., $\neg\exists G\big(\Diamond(E!G \;\&\; \mathcal{S}G) \;\&\; G = U\big)$, follows from the claim that unicorns are a mythical species, i.e., $Fictional\text{-}\mathcal{S}(U)$, by (.16).

Moreover, an even more general claim in Kripke 1972 [1980] can be derived from the assumption that unicorns are a mythical species, namely, that *there couldn't have been unicorns*. We may analyze this claim as $Fictional\text{-}S(U) \rightarrow \neg\Diamond\exists x Ux$. This is an instance of the following, more general theorem, in which x is a variable of type i, F is a variable of type $\langle i\rangle$, and R is a variable of type $\langle\langle i\rangle\rangle$, namely (.17) if F is a fictional R, then it is not possible that there is an individual x that exemplifies F:

(.17) $Fictional\text{-}R(F) \rightarrow \neg\Diamond\exists x Fx$

So if we instantiate F in (.17) to the (abstract) property *being a unicorn* (U), and instantiate R to the higher-order property *being a species* (S), we have derived the following strong, modal conclusion in Kripke 1972 [1980, 24]:

> ... it is said that though we have all found out that there are no unicorns, of course there might have been unicorns. Under certain circumstances there would have been unicorns. And this is an example of something I think is not the case.

This strong, modal conclusion, that it is not the case that there might have been unicorns (i.e., there couldn't have been unicorns), is put forward in a context in which Kripke is talking about the fact that some properties are contingently empty. But unlike the contingently empty property *giraffe in the Arctic Circle*, Kripke is *anticipating* the view he develops later (and that we represented above), namely, that unicorns are a mythical species and thus can't be identified with any possible species. Thus, *being a unicorn* is not just contingently empty, but necessarily empty. This is a consequence of our definitions and theorems, if given the assumption that *being a unicorn* is fictional.

The final issue to discuss, when considering how to type the principles that apply to fiction, concerns story identity. Philosophers are rightyly fond of stories that play around with the notion of identity. In many such cases, it may be that a *fictional relation of identity* is involved. Thus, when we import and analyze the identity claims that are made in some story $\underline{s}$, we may have to index the symbol '=' to $\underline{s}$ and regard the relation in question as an original character of the story.

Our system makes sense of this technique. If the notion of identity in question is one that relates individuals, then we may regard $=_{\underline{s}}$ as a symbol of type $\langle i, i\rangle$. But more generally, for every story $\underline{s}$ in which identity plays a role among the objects at some type level t (e.g., the story attributes certain, possibly odd, properties to this relation, or either says or relevantly implies that certain characters of type t are so related), we may:

- take $=_{\underline{s}}$ to be a relation having type $\langle t, t\rangle$,
- assert $OriginalCharacter(=_{\underline{s}}, \underline{s})$,
- take $\underline{s}$ to be attributing certain (odd) higher-order properties to this relation, and
- take $\underline{s}$ to be relating certain the characters of type t as via $=_{\underline{s}}$.

It would then be straightforward to apply (.9) and identify $=_s$ as the abstract relation of type $\langle t, t\rangle$ that encodes exactly the properties of such relations that hold of $=_{\underline{s}}$ in $\underline{s}$; i.e., where x has type $\langle t, t\rangle$, and both $A!$ and F have type $\langle\langle t, t\rangle\rangle$, assert:

(.18) $OriginalCharacter(=_{\underline{s}}, \underline{s}) \rightarrow \big(=_{\underline{s}} = \imath x(A!x \,\&\, \forall F(xF \equiv \underline{s} \models F{=_{\underline{s}}}))\big)$

With such a relation we may refine the notion of a character of a story given in (.6) above, so that the definition becomes *x is a character of* $\underline{s}$ just in case it is true in $\underline{s}$ that for every type-t-object y that is distinct$_{\underline{s}}$ from x, x exemplifies some property that that y fails to exemplify:

(.19) $CharacterOf(x,\underline{s}) \equiv_{df} \underline{s} \models \forall y(y \neq_{\underline{s}} x \rightarrow \exists F(Fx \,\&\, \neg Fy))$

The foregoing definitions may be useful if there are stories where there are distinct, original characters that are indiscernible from the point of view of the story.[404]

15.6 Analysis of Theoretical Mathematics

We turn now to theoretical mathematics, as distinguished from natural mathematics. In earlier chapters, we've analyzed a variety of natural mathematical objects, such as natural classes (logical sets), directions, shapes, and numbers. Indeed, we've derived, from just the primitive notions and axioms of object theory, some important parts of mathematics, namely, a number of definitions and principles of set theory, the axioms for a Boolean algebra (of concepts), and all of second-order Peano Arithmetic, including the existence of $\aleph_0$.

But now we analyze of the language of theoretical mathematics, i.e., mathematical theories formulated with primitive mathematical notions.[405] For example, (2nd-order) Zermelo-Fraenkel (ZF) set theory will be among our targets, and to formulate that theory, one needs only the primitive $\in$ (membership) as the sole non-logical 2-place relation constant. Indeed, notwithstanding the work in the previous chapter, (2nd-order) Peano Arithmetic (PA) can also serve as a target, by treating it as a theory that axiomatizes the following

[404]It is not clear to me that we need this more sophisticated definition of *character of* to handle the Frackworld case in Everett 2005. Everett says "I think it is pretty clear that in this story, it is left indeterminate as to whether Frick is Frack" (2005, 629). He uses a principle (P2) to infer from this that it is indeterminate as to whether the character Frick is the same object as the character Frack. This is allegedly problematic because of an argument by Evans (1978) in which the expression "being indeterminately identical to x" plays a key role, as if it were established that this expression definitively named a property. It isn't clear to me what theory of properties grounds such a claim. But even if the expression denotes a well-defined property, I don't think one can draw any conclusions from the fact that neither $s \models Frick =_{\underline{s}} Frack$ nor $s \models Frick \neq_{\underline{s}} Frack$. Stories are incomplete, and to analyze the particular Frackworld story, one may simply suppose that the story relevantly implies that Frick exemplifies the properties $[\lambda x \, \Diamond x =_{\underline{s}} Frack]$ and $[\lambda x \, \Diamond x \neq_{\underline{s}} Frack]$, but not $[\lambda x \, \Diamond x \neq_{\underline{s}} Frick]$. β-Conversion may be a principal of the relevant entailment and so hold within the scope of the story operator, but it isn't clear that the necessity of identity$_{\underline{s}}$ is also such a principle.

[405]I'm greatly indebted to Hannes Leitgeb and Uri Nodelman for our weekly discussions during which we composed Leitgeb, Nodelman, & Zalta m.s. I learned a tremendous amount from working with them and our joint research has had a profound impact on the exposition in this chapter.

mathematical *primitives*: the individual constant '0' (Zero), a 1-place relation constant 'N' (being a number), and the 2-place relation constant 'S' (successor). We'll then take PA to be the body of theorems derivable from Dedekind-Peano axioms, the axioms for recursive addition and multiplication, the principle of mathematical induction, and the comprehension principle for properties of numbers. The purpose of treating PA separately as part of theoretical mathematics is simply to give a simple example of well-known number theory that can be analyzed using the techniques introduced below.

15.6.1 Mathematical Theories

(**979**) **Remark:** The Normalization Methods for Formulating Mathematical Theories. We start with some simplifying assumptions about mathematical theories. These assumptions jointly constitute methods for normalizing and consistently formulating mathematical theories. Let T be a variable of our metalanguage that ranges over objects that we intuitively judge, on the basis of mathematical practice, to be mathematical theories. In our analysis of mathematics, we shall make use of the following assumptions, some of which will be discussed in more detail below:

(.1) each mathematical theory T is distinguished by a body of *closed* theorems formulable in terms of at least one distinguished, primitive mathematical term (i.e., at least one distinguished relation constant, and zero or more distinguished individual constants); we write $\vdash_T \varphi$ whenever φ is a theorem of T (and $\Gamma \vdash_T \varphi$ whenever φ is derivable from Γ in T);

(.2) the theorems of T consist of the mathematical axioms of T and their deductive consequences, though we allow for theories that are not formulated axiomatically;

(.3) each mathematical theory T has been formulated in the higher-order logic expressible in the language of relational type theory, without identity but with *closed* higher-order n-ary λ-expressions ($n \geq 1$) that are governed by higher-order principles of α-, β-, and η-Conversion (see the observation below for explanation);

(.4) each n-ary function term τ of T has been replaced by an appropriate n+1-ary relation term Π, and the axioms governing τ have been reformulated as equivalent axioms governing Π,[406] and (c) any formulas containing

[406]See, for example, treatments of function terms within the predicate calculus with equality in Mendelson 1964 [1997] and Enderton 1972 [2001]. In Chapter 2, Section 9 of Mendelson 1964 [1997, 103ff], we find:

> In mathematics, once we have proved, for any $y_1,\dots,y_n$, the existence of a unique

definite descriptions have been replaced by equivalent formulas that assert existence and uniqueness claims;

(.5) the language and theorems of T may include non-primitive (i.e., defined) terms τ as long as τ is both *closed* and *uniquely-* or *well-defined* in T; i.e., we may assume that T includes theorems with the defined term τ as long as τ has been added to language of T and axiomatized by taking the definition of τ as an additional axiom or principle;

(.6) mathematical theories T and T' that are notational variants may be identified;

(.7) axiomatic theories T and T' may be identified whenever T' has a redundant axiom, since T and T' have the same theorems;

(.8) whenever T is formulated with an expression of some type t other than $\langle\,\rangle$,[407] then for each such type t, the theory T includes identity^t ('$=^t$') as a distinguished relation symbol having type $\langle t, t\rangle$ governed by classical axioms and rules, unless $=^t$ is specially systematized otherwise by T (henceforth, we suppress the superscript t in $=^t$ since this can be inferred from the types of the terms flanking the $=$ symbol); and

object u that has the property $\mathcal{B}(u, y_1, \ldots, y_n)$, we often introduce a new function letter $f(y_1, \ldots, y_n)$ such that $\mathcal{B}(f(y_1, \ldots, y_n), y_1, \ldots, y_n)$ holds for all $y_1, \ldots, y_n$. In cases where we have proved the existence of a unique object u that satisfies a wf $\mathcal{B}(u)$ and $\mathcal{B}(u)$ contains u as its only free variable, then we introduce a new individual constant b such that $\mathcal{B}(b)$ holds. It is generally acknowledged that such definitions, though convenient, add nothing really new to the theory.

Mendelson then makes this claim precise with Proposition 2.29, where we find a proof that any theory T with well-defined function terms can be reformulated in terms of a theory T' without such terms, via a mapping # for which it can be shown that if $\vdash_T \varphi$, then $\vdash_{T'} \varphi^{\#}$. And he sets an exercise (Exercise 2.83) asking the reader to transform formulas with function terms to formulas without.

Enderton 1972 [2001] (§2.7) establishes both model-theoretic and proof-theoretic results underlying this phenomenon. His remarks have a direct bearing on the examples we construct below. See, for example, 1972 [2001, 165], where he considers extending a theory T without function symbols to a theory with the *unary* function symbol f by introducing f in to the language with the definition:

(δ) $\forall x \forall y (f(x) = y \equiv \varphi)$

He then shows (1972 [2001, Theorem 27A]) that $\forall x \exists! y \varphi$ is a theorem of T if and only if, for any sentence ψ in the language without f, if ψ is valid in any model of T+(δ), then ψ is valid in T. Just as importantly for our purposes, the discussion (1972 [2001, 169–172]) gives a translation procedure for eliminating the function terms generally, i.e., for mapping formulas with n-ary function terms to formulas in which n+1-ary relation terms replace the function terms. We'll see examples like this below. Enderton establishes (1972 [2001, Theorems 27B–27D]) that the procedure achieves the goal of eliminating the function terms without loss.

[407] I.e., whenever T is formulated with an expression for an individual (of type $t = i$) or an expression for an non-zero n-place relation having type $t = \langle t_1, \ldots, t_n\rangle$ for some types $t_1, \ldots, t_n$ ($n \neq 0$)

(.9) the axioms and rules of the logic L assumed by a mathematical theory T have been included among the axioms and rules of T, so that $\vdash_T \varphi$ is to be strictly regarded as shorthand for $\vdash_{T_L} \varphi$; when T is formulated in classical (higher-order) logic, we may use the simpler notation $\vdash_T \varphi$ to indicate theoremhood in T, but if T is formulated within a *non-classical* (higher-order) logic L, then it becomes informative to reference the theorems of this theory by writing $\vdash_{T_L} \varphi$.

We take these assumptions to be relatively uncontroversial and consistent with mathematical practice. One could attempt to give a deeper level of analysis of mathematical theories within object theory by starting with pre-theoretic judgments about which propositions are *mathematical*, but we shall not officially make the attempt here.[408]

However, some discussion of assumptions (.3), (.5), (.8), and (.9) may prove to be helpful. In connection with (.3), we first discuss an example and then provide some justification. As an example, suppose a theory T includes, as theorem, the simple relational statement Rab, where a and b are constants of type i and R a constant of type $\langle i, i\rangle$. Given assumption (.3), the language of T includes the closed terms $[\lambda x\, Rxb]$, $[\lambda x\, Rax]$, and $[\lambda F\, Fab]$ (the types of which are obvious), and T itself includes the following instances of β-Conversion as theorems (indeed, if T is formulated axiomatically, the following would be axioms, since typically, mathematical theories don't assume free logic):

- $[\lambda x\, Rxb]a \equiv Rab$
- $[\lambda x\, Rax]b \equiv Rab$
- $[\lambda F\, Fab]R \equiv Rab$

The last asserts that R exemplifies the higher-order property *being a binary relation F that a and b exemplify* if and only if a and b exemplify R.

[408] Unofficially, we might take $Mathematical(p)$ as primitive and then define a *mathematical situation* to be any situation s such that every proposition true in s is a mathematical proposition, i.e., where s is a variable ranging over situations:

$$MathSituation(s) \ \equiv_{df} \forall p((s \models p) \rightarrow Mathematical(p))$$

Then, using the notion of *authorship* from the theory of fiction, one could define a *mathematical theory* as a mathematical situation that some concrete object authored:

$$MathTheory(x) \ \equiv_{df} MathSituation(x) \,\&\, \exists y(E!y \,\&\, Ayx)$$

Were such a definition available in the present theory, we could introduce the variable T in the object language, to range over the defined notion of a mathematical theory. This defined notion could help to explain how mathematical practice changes the expressive power of language in contingent ways (e.g., by the introduction of new expressions into the language). This may be important when we consider mathematics and modality. Note, however, that one could define a *possible mathematical theory* as a mathematical situation such that it is possible that some concrete authored it.

Since Rab is a theorem, it follows that the theorems of T include: $[\lambda x\, Rxb]a$, $[\lambda x\, Rax]b$, and $[\lambda F\, Fab]R$, $[\lambda xy\, Rxy]ab$, and $[\lambda Fxy\, Fxy]Rab$. Of course, the presence of other closed, higher-order λ-expressions also yields additional facts about T, but these will be analyzed along the same lines as the analysis of the ones just discussed.[409] These facts play a key role when we identify the mathematical relation R as an abstract relation that encodes the properties of relations that are attributed to R in T.

We may justify the formulation of mathematical theories with closed higher-order λ-expressions by examining a more complex example. Let T_1 be the classical theory of dense, linear orderings without endpoints, axiomatized as follows, in which $<$ and $=$ are the only primitive relation terms:

$$\forall x \forall y \forall z(x < y \,\&\, y < z \rightarrow x < z) \qquad \text{(Transitivity)}$$
$$\forall x \neg(x < x) \qquad \text{(Irreflexivity)}$$
$$\forall x \forall y(x \neq y \rightarrow (x < y \vee y < x)) \qquad \text{(Connectedness)}$$
$$\forall x \forall y \exists z(x < z < y) \qquad \text{(Dense)}$$
$$\forall x \exists y \exists z(z < x < y) \qquad \text{(No Endpoints)}$$

Clearly, it is consistent with and, indeed, a part of, mathematical practice to talk about the relation $<$ axiomatized by T_1. For example, one may talk about the properties $<$ has in T_1, for the topic "properties of relations" is a well-known topic. We may say, for example, that $<$ has the property of being irreflexive and transitive. Our closed λ-expressions allow us to formally express this understanding by noting that the following are implied by, and so among the theorems of, T_1, by way of higher-order β-Conversion:

$[\lambda F\, \forall x \forall y \forall z(Fxy \,\&\, Fyz \rightarrow Fxz)]<$
i.e., $<$ exemplifies *being a transitive binary relation*

$[\lambda F\, \forall x(\neg Fxx)]<$
i.e., $<$ exemplifies *being an irreflexive binary relation*

[409]So, for example, the language of T also includes the closed terms $[\lambda xy\, Rxy]$ and $[\lambda Fxy\, Fxy]$, and so T also has the following instances of β-Conversion as theorems:

- $[\lambda xy\, Rxy]ab \equiv Rab$
- $[\lambda Fxy\, Fxy]Rab \equiv Rab$

Thus, T also includes the theorems $[\lambda xy\, Rxy]ab$ and $[\lambda Fxy\, Fxy]Rab$. The λ-expressions in these theorems will be represented using terms that denote abstract relations of T.

Moreover, T also includes the following instance of η-Conversion (935.27.a) as a theorem: $[\lambda xy\, Rxy] = R$. In these cases, we will identify the relation $[\lambda xy\, Rxy]$ of T using the same methodology we develop below for assigning a denotation to the properties $[\lambda x\, Rxb]$, $[\lambda x\, Rax]$, and $[\lambda F\, Fab]$.

Finally, since $[\lambda G\, [\lambda F\, Fab]G]$ is a closed λ-expression in the language, T will have the following instance of η-Conversion as a theorem: $[\lambda G\, [\lambda F\, Fab]G] = [\lambda F\, Fab]$. Again, our methodology will identify the abstract relations denoted by both λ-expressions flanking the identity sign.

Again, these facts play a key role when we identify the distinguished relation $<$ of T_1 as the abstract relation that encodes the properties of relations that are attributed to $<$ in T.

Before we move on to the next assumption in need of discussion, it is worth saying *why* we have not allowed the language of T to be extended with *open* λ-expressions. To see why, we have to set up a simple example. Take PA as a primitive theory and consider the closed theorem $\forall x(x{+}2 = 2{+}x)$. Since we're assuming that PA has been formulated without function terms, such a theorem would be expressed with a ternary relation symbol, say R, instead of the binary function symbol $+$, where $Rxyz$ asserts: x added to y yields z. Moreover, R would appropriately axiomatized so as to capture the recursive axioms for addition.[410] Thus, the theorem $\forall x(x{+}2 = 2{+}x)$ could be expressed relationally as $\forall x\forall z(Rx2z \equiv R2xz)$. Now suppose we had allowed PA to be formulated with open λ-expressions. Then, by β-Conversion and the fact that $\forall z(Rx2z \equiv R2xz)$ is equivalent to $[\lambda y\, \forall z(Rxyz \equiv Ryxz)]2$, we would be able to use the Rule of Substitution to derive the following as a theorem of PA:

$$\forall x([\lambda y\, \forall z(Rxyz \equiv Ryxz)]2)$$

(This is easier to see with function terms: the strict equivalence of $x{+}2 = 2{+}x$ and $[\lambda y\; x{+}y = y{+}x]2$ would make $\forall x([\lambda y\, \forall x(x{+}y = y{+}x)]2)$ a consequence of $\forall x(x{+}2 = 2{+}x)$.) But such a theorem might lead one to ask the question, "What is the denotation of the (open) expression $[\lambda y\, \forall z(Rxyz \equiv Ryxz)]$ as it occurs in the language of PA?" In what follows, we shall not accept this as a valid question, for this open λ-expression has a different denotation for each value of x. Asking such question is analogous to asking the question "What is the denotation of the free variable x in the language of PA?". Free variables, and complex

[410] In other words, we can restate the axioms for recursive addition without the function terms by applying Enderton's algorithm (1972 [2001, 169]) as follows. We start with a ternary relation term R for which it is stipulated that two conditions hold. To formulate these two conditions, we first turn the axioms for recursive $+$, which are stated with the help of the successor function term x', into axioms that replace the binary function $+$ by R. So we turn:

$$\forall x(x + 0 = x)$$

$$\forall x\forall y(x + y' = (x + y)')$$

into the following relational axioms:

$$\forall x(Rx0x)$$

$$\forall x\forall y\forall z(Rxyz \to Rxy'z')$$

Then, we eliminate the function terms y' and z' by using the primitive *predecessor* relation P, for which $\forall y\exists!xPyx$ holds, by introducing the definition $\forall y\forall x(y' = x \equiv Pyx)$ (à la Enderton 1972 [2001, 165]), so that the second axiom above becomes via Enderton's algorithm:

$$\forall x\forall y\forall z\forall u\forall v(Rxyz \to (Pyu \to (Pzv \to Rxuv)))$$

This asserts the universal generalization of: if x added to y yields z, y precedes u, and z precedes v, then x added to u yields v. Note that by defining R in this way, it becomes provable that R is functional, i.e., that $\forall x\forall y\exists!zRxyz$.

mathematical terms containing free variables (i.e., 'complex variables'), do not have denotations *simpliciter*, and so don't induce the same kind of ontological question as "What is the denotation of the term 0 in PA?" or "What is the denotation of the term ∅ in ZF?" (Or, rather, we might say: answers to the ontological questions about the denotation of closed terms have to be *in hand* if we are to advance from the semantic analysis of variables as expressions that can take different values to the ontological question of what those values are.) Thus, the questions about the denotation of the expressions $[\lambda y\, \forall z(Rxyz \equiv Ryxz)]$ and x are analogously ill-conceived since both are variables: x is a simple variable and the λ-expression is a complex variable.

By constrast, we shall answer questions of the form "What is the denotation of $[\lambda y\, \forall x \forall z(Rxyz \equiv Ryxz)]$, where this is a closed λ-expression. That is, from the theorem $\forall x \forall z(Rx2z \equiv R2xz)$ with which we started, we will also take $[\lambda y\, \forall x \forall z(Rxyz \equiv Ryxz)]2$ to be a theorem of PA. And it will be a theorem that $[\lambda F\, \forall x \forall z(Fx2z \equiv F2xz)]R$. The λ-expressions involved in these theorems pose ontological questions and we shall, in what follows, interpret the first λ-expression as having type $\langle i \rangle$ and as denoting an abstract property of individuals, and the second as having type $\langle\langle i,i,i\rangle\rangle$ and as denoting an abstract property of ternary relations among individuals.

Now consider (.5). Though we've assumed in (.4) that mathematical theories are formulated without function terms, (.5) relaxes this assumption somewhat, since it allows us to consider examples of mathematical theorems expressed in defined notation rather than primitive notation, provided those theorems are expressed with defined constants and closed function terms. Even a simple theorem of ZF such as $\varnothing \in \{\varnothing\}$ involves the defined constant $\varnothing$ and the defined, closed function term $\{\varnothing\}$. To bring these theorems into the analytic fold of object theory, we may regard the language of ZF as extended with the primitive symbol $\varnothing$ and the primitive term $\{\varnothing\}$, where these primitive terms are governed by axioms that simply reformulate their usual definitions. Thus, for $\varnothing$, we may suppose that ZF is extended with the closures of the axiom:

(ϑ) $\varnothing = x \equiv \forall z(z \in x \equiv z \neq z)$

This guarantees that $\varnothing$ is a well-defined term, since it is a theorem of ZF that there is a unique x that has as members anything that fails to be non-self-identical, i.e., $\exists! x \forall z(z \in x \equiv z \neq z)$. So a unique set satisfies the left condition of (ϑ).

And for $\{\varnothing\}$, we may suppose ZF is extended with the closures of the axiom:

(ξ) $\{\varnothing\} = x \equiv \forall z(z \in x \equiv z = \varnothing)$

This ensures that $\{\varnothing\}$ is a well-defined term, since it is a theorem of ZF that there is a *unique* x such that x has y and only $\varnothing$ as a member, i.e., $\exists! x \forall z(z \in x \equiv z = \varnothing)$. So a unique set satisfies the left condition of (ξ). Together, (ϑ), (ξ), (a), and (b) ensure that any theorem in ZF about $\varnothing$ and $\{\varnothing\}$ (e.g., $\varnothing \in \{\varnothing\}$) is equivalent to a theorem expressed entirely in primitive notation without these symbols, just as if they had been introduced by definition.

As another example from set theory, consider that where a is some named set, then $\cup a$ ('the union of a') may be defined by:

$$\cup a = x \equiv_{df} \forall y(y \in x \equiv \exists z(z \in a \,\&\, y \in z))$$

$\cup a$ is well-defined because the Extensionality and Union axioms of ZF combine to imply, for any well-defined set a:

$$\exists! x \forall y(y \in x \equiv \exists z(z \in a \,\&\, y \in z))$$

Thus, we may, in what follows, consider theorems of ZF containing the term $\cup\kappa$, for any well-defined, closed term κ in ZF. As an example, consider $\cup\{\varnothing, \{\varnothing\}\}$.

More generally, we may consider data from theoretical mathematics expressed with well-defined, closed terms such as $4!$, 2^8, $\sqrt{2}$, $\Sigma_{n=1}^{5} n{+}3$, π, ω, etc., provided we import their definitions. We need not recast theorems containing such closed function terms as theorems without such terms. However, for the same reasons discussed in connection with open λ-expressions in (.3) above, we won't consider data with open function terms.

Thus, for convenience, (a) we shall not suppose that PA includes open function terms such as $n!$, x^2, 2^x, $\sqrt{y}$, etc., or that ZF includes the open function terms such as $x \cap y$, $\{x\}$, etc., and (b) we shall not consider such facts as:

(ζ) $\vdash_{\text{PA}} \exists x(x^2 = 4)$

(ξ) $\vdash_{\text{ZF}} \exists x(\varnothing \in \{x\})$

If were to treat (ζ) and (ξ) as data, then the question 'What do the terms denote?' applies only to the symbols $=$ and 4 in (ζ) and to the symbols $\varnothing$ and $\in$ in (ξ). We shall not accept the question "What is denoted by the open terms x^2 and $\{x\}$?" This question is ill-conceived because x^2 and $\{x\}$ are complex variables and take on a different denotation for each value of x. Consequently, just like simple variables, they doesn't pose an ontological problem, or rather, a semantics of the simple variables of a theory T assumes we have already solved the ontological problem of what those variables range over. An ontological problem only arises for closed terms like 2^2 and $\{\varnothing\}$.[411]

[411] As described in footnotes 406 and 410, the loss of function terms with free variables is not onerous. Consider the case of the unit set of x, i.e., $\{x\}$. In mathematical practice, one might introduce this term via the definition:

$$\{x\} = y =_{df} \forall z(z \in y \equiv z = x)$$

or by adding primitive expressions of the form $\{x\}$ for any variable x, governed by the axiom:

$$\forall(\{x\} = y \equiv \forall z(z \in y \equiv z = x))$$

In connection with (.8) above, it should be observed that mathematical theories are traditionally formulated only in terms of expressions for individuals and (possibly higher-order) n-place relations ($n \neq 0$); they are not formulated with primitive constants or terms for 0-place relation. Mathematical practice doesn't give any evidence that mathematicians are interested in propositions *per se*, or in the identity of propositions. In other words, mathematics appears to be the study of abstract individuals and abstract n-place relations, for $n \neq 0$, and the propositions of mathematics are *ordinary* propositions whose constituents are abstract.

Finally, in connection with (.9), consider Heyting Arithmetic (HA), which uses the same language and non-logical axioms as PA but asserts the latter in the context of intuitionistic predicate logic (IQC). Though we can regard HA as a single deductive system comprising the logical axioms and rules of IQC and the non-logical axioms of PA, it is informative to regard the proof-theoretic claim $\vdash_{\mathrm{HA}} \varphi$ as having the form $\vdash_{T_L} \varphi$, where T = PA and L = IQC So the claim $\vdash_{\mathrm{HA}} \varphi$ becomes a claim of the form $\vdash_{T_L} \varphi$. But either way, we can use the methods outlined above to target (the theorems of) HA for the analysis described below.

(**980**) **Axiom, Definition, and Theorems:** Extending Object Theory to Analyze *Truth In*, and *Identity*, for Mathematical Theories. Since we've taken mathematical theories to be distinguished by their theorems, we shall analyze them as situations, i.e., objects that encode only propositions (and in particular, those propositions that are their theorems). To state this analysis, we extend the language of object theory with new constants that (intuitively) name mathematical theories, such as ZF, PA, $\mathbb{R}$ (real number theory), $\mathbb{C}$ (complex number theory), $< D$ (dense linear orderings without endpoints), etc. Where κ is any new constant introduced as a name of a mathematical theory, we assert, as an axiom, that:

(.1) $Situation(\kappa)$ (**Axiom**)

Moreover, we engage in a harmless abuse of notation: the expression 'T' has, up to now, been used intuitively as a variable of our metalanguage ranging over what we pretheoretically judge to be mathematical theories. However we

But we may logically 'rehabilitate' this nicety of mathematical practice by instead extending the language of ZF with a binary relation constant, say R, governed by the principle:

$$\forall x \forall y (Rxy \equiv x \in y \,\&\, \forall z (z \in y \rightarrow z = x))$$

I.e., x and y exemplify R just in case x is a unique member of y. Since ZF, in a higher-order setting, implies that there is such a relation and that it is functional (i.e., $\forall x \exists ! y Rxy$), we can apply the the standard procedures for identifying a theorem that expresses what (ξ) expresses, namely, $\exists x \forall y (Rxy \rightarrow \varnothing \in y)$. In this latter theorem, we only have to find denotations for the constants 'R', '$\varnothing$', and '$\in$', and give truth conditions to the theorem as a whole.

shall now use 'T' technically as a variable of type i ranging over those situations introduced as mathematical theories. For a justification of this abuse of notation and a way to eliminate the abuse, see the suggestions described in footnote 408. However, if one is uncomfortable introducing the variable T in this way, simply restate the principles below that use the variable T as a schema in which the metavariable κ (which ranges over those new constants introduced to name mathematical theories) replaces the variable T.

Now given the type-theoretic definition of a situation in (971.2), we may then appeal to the defined notion *p is true in situation s* (971.3) to say what it is for a proposition p to be *true in T*:

(.2) p *is true in* $T =_{df} T \models p$

Hence, by (971.3) and (971.1), p is true in T just in case $T[\lambda y\, p]$, where y is a variable of type i bound vacuously in $[\lambda y\, p]$.

Since a mathematical theory T is a situation, it now follows that (.3) T is identical to *the* abstract individual that encodes a property of individuals F iff F is a propositional property of the form $[\lambda y\, p]$ for some proposition true in T:

(.3) $T = \imath x(A!x \,\&\, \forall F(xF \equiv \exists p(T \models p \,\&\, F = [\lambda y\, p])))$ (**Theorem**)

We may simplify the statement of this theorem by appealing to canonical situation descriptions, i.e., T is the situation s that makes true all and only the propositions true in T:

(.4) $T = \imath s \forall p((s \models p) \equiv (T \models p))$ (**Theorem**)

It is to be emphasized that this is not a *definition* of T, but rather a principle by which we can theoretically identify T in terms of data of the form $T \models p$. For example, we may theoretically identify ZF as follows:

- $\mathrm{ZF} = \imath s \forall p((s \models p) \equiv (\mathrm{ZF} \models p))$

Thus, ZF is the situation (i.e., the abstract individual) that makes true exactly the propositions true in ZF.

Of course, for these principles to be informative, we have to have data, expressible in object theory, of the form $\mathrm{ZF} \models \varphi$ and, more generally, of the form $T \models \varphi$. We accomplish this with the Importation Principle, formulated in (981.3) below. But first notice that this understanding of mathematical theories allows us to distinguish a theory from its axiomatization. The same theory (i.e., the same body of theorems) can be axiomatized in different ways, but we identify the theory with the (abstract object that encodes all and only the) theorems, rather than with some particular axiomatization of it.

(**981**) **Definitions, New Terms, Meta-axioms and Metarules:** The Importation Principle and Rule of Closure. We now work our way to the statement of a principle that will introduce axioms that are analytic truths. Whenever φ is a theorem of T, the principle asserts that a notational variant of φ, labeled φ^*, is true in T. Thus, the principle asserts axioms of the form $T \models \varphi^*$ for each mathematical theory T. To state this principle, we we must first (a) define a metalinguistic, syntactic notion needed to state this principle, and (b) extend our language with some new expressions that are used in the notational variant φ^*.

The notion we need to define is that of a *primary term of* φ. This notion has already been defined for exemplification and encoding formulas, in (930) [7]. But we now extend the notion so that we may designate the primary terms of an arbitrary formula φ. We define this metalinguistic notion by recursion as follows:

(.1) τ is a *primary term of* φ if and only if either

- φ is an exemplification formula $\Pi\tau_1 \ldots \tau_n$ $(n \geq 1)$ and τ is one of Π, $\tau_1, \ldots, \tau_n$, or
- φ is an encoding formula $\tau_1 \ldots \tau_n\Pi$ $(n \geq 1)$ and τ is one of Π, $\tau_1, \ldots, \tau_n$, or
- φ is any other formula and τ is a primary term in a proper subformula of φ.

Note the following consequences of this definition:

- No formula is a primary term of itself or of any other formula.
- If φ is a formula in virtue of being a constant, a variable, or a description of type $\langle\,\rangle$, then φ has no primary terms, since such a φ is neither an exemplification formula, nor an encoding formula, nor has any proper subformulas. Moreover, such formulas contribute no primary terms to any formula of which they may be a subformula.
- If a λ-expression $[\lambda\alpha_1 \ldots \alpha_n\, \psi]$ or a description $\iota\alpha\psi$ occur in a formula φ, then none of the terms of ψ are primary terms of φ, since ψ doesn't qualify as a subformula of φ.

Some examples of the foregoing are:

- The constant p_1 and variable p, both of type $\langle\,\rangle$, have no primary terms, and neither does any description of the form $\iota p\varphi$.
- The primary terms of $Fx \rightarrow p$ and $Fx \rightarrow \iota p(p \,\&\, \neg p)$ are just F and x.
- The primary terms of the formula $[\lambda x\, \neg Fx \,\&\, Pb]a$ are a and $[\lambda x\, \neg Fx \,\&\, Pb]$; the primary terms of the formulas $\neg Fx$, Fx, Pb, and $\neg Fx \,\&\, Pb$ (i.e., the primary terms of the subformulas of the matrix $\neg Fx \& Pb$) are not primary terms of $[\lambda x\, \neg Fx \,\&\, Pb]a$.

- The primary terms of of the formula $P\imath xRxb$ are just P and $\imath xRxb$; the primary terms of the matrix of the description are not primary terms of $P\imath xRxb$.

Moreover, if we consider mathematical theories formulated in accordance with (979.1) – (979.9), then we have the following examples:

- If PA is extended with the constants '2' and '4', and the closed function term '2^2', then the primary terms of the theorem $2^2 = 4$ are just '2^2', '=', and '4'. The numeral '2' is not one of the primary terms of $2^2 = 4$, though it will occur as a primary term of other theorems of PA.

- If ZF is extended with the constant '$\varnothing$' and the closed function term '$\{\varnothing\}$', then the primary terms of the theorem $\{\varnothing\} \in \{\{\varnothing\}\}$ are '$\{\varnothing\}$', '$\in$', and '$\{\{\varnothing\}\}$'. The symbol '$\varnothing$' is not one of the primary terms of $\{\varnothing\} \in \{\{\varnothing\}\}$, though it will occur as a primary term of other theorems of ZF.

Now we may extend our language with new terms that have been relativized to a mathematical theory, as follows.

Let T be any mathematical theory formulated in accordance with (979.1) – (979.9). Then for any formula φ such that $\vdash_T \varphi$, the *closed*, primary terms of φ are the primary terms of φ other than a variable. (Since we've already eliminated the open complex terms of T, such as open λ-expressions and open function terms, the only remaining open terms of T are the variables; so even if a variable α appears as a primary term of φ, α is not a closed primary term of φ.) We then extend our language by stipulating:

(.2) Whenever $\vdash_T \varphi$, and τ is a closed primary term of φ, the expression τ_T shall be a new (closed) term having the same type as τ.

Thus, depending on the formulation of T, we may introduce new terms of the form τ_T, where τ is either (a) a constant of type i or of type $\langle t_1,\ldots,t_n\rangle$ $(n \geq 1)$, (b) a closed n-ary function term of the form $f(\tau_1,\ldots,\tau_n)$ $(n \geq 1)$, where $\tau_1,\ldots,\tau_n$ have types $t_1,\ldots,t_n$, respectively, or (c) a closed λ-expression of the form $[\lambda\alpha_1\ldots\alpha_n\,\varphi]$ $(n \geq 1)$, where $\alpha_1,\ldots,\alpha_n$ have types $t_1,\ldots,t_n$, respectively. We shall *not* be adding any of the following: indexed variables, indexed open function terms, indexed open λ-expressions, or indexed formulas.

With these stipulations we may now formulate and assert the Importation Principle as a meta-axiom, i.e., a principle that describes conditions under which new axioms of a certain form are to be added to our system:

(.3) **Importation Principle**. Whenever $\vdash_T \varphi$ and φ^* is the formula of object theory that results by substituting τ_T for each closed primary term τ in φ, then $T \models \varphi^*$ is an (analytically true) axiom.

Given that we've extended our language with terms of the form τ_T, the formula $T \models \varphi^*$ is an expression in the language of our system. Where 'S' ('*being a set*') is a ZF relation constant of type $\langle i \rangle$ and the other symbols (and their types) are obvious, here are some examples of importation pairs $\vdash_T \varphi$ and $T \models \varphi^*$, in which the second member of each pair may be regarded as axiomatic:

(.4) **Examples of Importation:**

(.a) $\vdash_{\text{ZF}} \varnothing \in \{\varnothing\}$
$\text{ZF} \models \varnothing_{\text{ZF}} \in_{\text{ZF}} \{\varnothing\}_{\text{ZF}}$

(.b) $\vdash_{\text{ZF}} \{\varnothing\} \in \{\{\varnothing\}\}$
$\text{ZF} \models \{\varnothing\}_{\text{ZF}} \in_{\text{ZF}} \{\{\varnothing\}\}_{\text{ZF}}$

(.c) $\vdash_{\text{ZF}} [\lambda x\, x \in \{\varnothing\}]\varnothing$
$\text{ZF} \models [\lambda x\, x \in \{\varnothing\}]_{\text{ZF}}\varnothing_{\text{ZF}}$

(.d) $\vdash_{\text{ZF}} [\lambda x\, \varnothing \in x]\{\varnothing\}$
$\text{ZF} \models [\lambda x\, \varnothing \in x]_{\text{ZF}}\{\varnothing\}_{\text{ZF}}$

(.e) $\vdash_{\text{ZF}} [\lambda F\, \varnothing F\{\varnothing\}]\in$
$\text{ZF} \models [\lambda F\, \varnothing F\{\varnothing\}]_{\text{ZF}}\in_{\text{ZF}}$

(.f) $\vdash_{\text{ZF}} \exists x(x \in \{\varnothing\})$
$\text{ZF} \models \exists x(x \in_{\text{ZF}} \{\varnothing\}_{\text{ZF}})$

(.g) $\vdash_{\text{ZF}} \neg\exists x(Sx \,\&\, x \in \varnothing)$
$\text{ZF} \models \neg\exists x(S_{\text{ZF}}x \,\&\, x \in_{\text{ZF}} \varnothing_{\text{ZF}})$

(.h) $\vdash_{\text{ZF}} [\lambda y\, \neg\exists x(Sx \,\&\, x \in y)]\varnothing$
$\text{ZF} \models [\lambda y\, \neg\exists x(Sx \,\&\, x \in y)]_{\text{ZF}}\varnothing_{\text{ZF}}$

(.i) $\vdash_{\text{ZF}} [\lambda F\, \neg\exists x(Fx \,\&\, x \in \varnothing)]S$
$\text{ZF} \models [\lambda F\, \neg\exists x(Fx \,\&\, x \in \varnothing)]_{\text{ZF}}S_{\text{ZF}}$

(.j) $\vdash_{\text{ZF}} [\lambda F\, \neg\exists x(Sx \,\&\, Fx\varnothing)]\in$
$\text{ZF} \models [\lambda F\, \neg\exists x(Sx \,\&\, Fx\varnothing)]_{\text{ZF}}\in_{\text{ZF}}$

(.k) $\vdash_{\text{PA}} \exists x(2{\times}2 = x)$
$\text{PA} \models \exists x((2{\times}2)_{\text{PA}} =_{\text{PA}} x)$

(.l) $\vdash_{\text{PA}} [\lambda F\, \exists x(2{\times}2Fx)]= \quad [F \text{ infix}]$
$\text{PA} \models [\lambda F\, \exists x(2{\times}2Fx)]_{\text{PA}}=_{\text{PA}}$

(.m) $\vdash_{\mathbb{R}} \forall x(x > 4 \to x > \pi)$
$\mathbb{R} \models \forall x(x >_{\mathbb{R}} 4_{\mathbb{R}} \to x >_{\mathbb{R}} \pi_{\mathbb{R}})$

(.n) $\vdash_{\mathbb{R}} [\lambda y\, \forall x(x > 4 \to x > y)]\pi$
$\mathbb{R} \models [\lambda y\, \forall x(x > 4 \to x > y)]_{\mathbb{R}}\pi_{\mathbb{R}}$

(.o) $\vdash_{\mathbb{R}} [\lambda yz\, \forall x(x > y \to x > z)]4\pi$
$\mathbb{R} \models [\lambda yz\, \forall x(x > y \to x > z]_{\mathbb{R}}4_{\mathbb{R}}\pi_{\mathbb{R}}$

(.p) $\vdash_{\mathbb{R}} [\lambda F\, \forall x(Fx4 \to Fx\pi)]>$
$\mathbb{R} \models [\lambda F\, \forall x(Fx4 \to Fx\pi)]_{\mathbb{R}}>_{\mathbb{R}}$

Now, as a consequence of our assumptions (979.2) and (979.9), we may derive two metarules. Using the φ^* notation introduced earlier, the first states that (.5) if φ is not a theorem of theory T, then it is not a theorem of our theory that $T \models \varphi^*$:

(.5) **Metarule:** If $\nvdash_T \varphi$, then $\nvdash T \models \varphi^*$

The proof in the Appendix appeals to a metalemma contributed by Uri Nodelman.

The second metarule, the Rule of Closure for Truth in a Theory, intuitively states that (.6) if ψ is a deductive consequence, in T, of formulas $\varphi_1, \ldots, \varphi_n$, and if $\varphi_1^*, \ldots, \varphi_n^*$ are all true in T, then ψ^* is true in T:

(.6) **Rule of Closure for Truth in a Theory**
If $\varphi_1, \ldots, \varphi_n \vdash_T \psi$ and $\vdash T \models \varphi_1^*$, and ... and $\vdash T \models \varphi_n^*$, then $\vdash T \models \psi^*$

It is important to remember that in this metarule, the claim $\varphi_1, \ldots, \varphi_n \vdash_T \psi$ is a claim about derivability among formulas in T. So the metarule tells us if a proof-theoretic fact about T and certain relevant deductive facts about T hold in object theory, then another deductive fact about T holds in object theory.

Thus, we may reason normally, in object theory, among the formulas φ such that $T \models \varphi$, though note that *we are limited to reasoning only among those formulas that result from importing theorems expressible in the language of* T.[412]

Now that we've extended our system by adding the second member of each pair as an axiom, our next goal is to identify the abstract objects denoted by the indexed terms of the form τ_T.

(**982**) **Remark:** Digression: Are the Imported Axioms Modally Fragile? I shall leave it as an open question as to whether the axioms introduced by the Importation Principle are modally fragile analytic truths. This is a deep question about which I am of two minds. One reason for supposing that axioms are modally fragile is that to instantiate our definitions and prove facts about mathematics, we have to be given data that arises in the course of mathematical practice. Mathematical practice involves contingent acts, utterances, and behaviors. Indeed, it is arguable that when mathematicians put forward a theory, they are changing the expressive power of the the language of mathematics. Putting forward a mathematical theory involves a special use of language. It is not a situation in which terms with an antecedently fixed meaning are used to make assertions, since the meanings of terms for abstract objects are not independent of our theory of those objects. Rather it is a situation in which the meanings of the terms are being introduced implicitly, by asserting axioms or a body of 'definitive' truths. There is an element of contingency to this, and to the extent that there is, it may be best to mark the analytic truths resulting from the Importation Principle as modally fragile. They will still be provably necessary, but (a) the proof depend on the necessity of encoding (935.30), rather than the derived Rule of Necessitation, and (b) the necessary truth that is established will accordingly be marked as having been derived from a modally fragile axiom.

On the other hand, it may be reasonable to ignore such contingencies in the present context. After all, we're analyzing mathematical language as it is

[412] Thus, from the fact *Socrates is wise or it is not the case that Socrates is wise* ($Ws \vee \neg Ws$) and the fact that $\mathrm{ZF} \models \varnothing_{\mathrm{ZF}} \in_{\mathrm{ZF}} \{\varnothing\}_{\mathrm{ZF}}$, the Rule of Closure for Truth in a Theory does *not* imply that $\mathrm{ZF} \models (Ws \vee \neg Ws)$, even though $Ws \vee \neg Ws$ is deductively implied by $\varnothing_{\mathrm{ZF}} \in_{\mathrm{ZF}} \{\varnothing\}_{\mathrm{ZF}}$. For it is not the case that $\varnothing \in \{\varnothing\} \vdash_{\mathrm{ZF}} (Ws \vee \neg Ws)$. $Ws \vee \neg Ws$ is not expressible in T.

now used and the history of that usage may not be relevant to the nature of the objects postulated by the theory. It might be argued that the imported analytic truths are necessarily (and always) true and that the *proof* of their necessity (and omnitemporality) stems solely from the fact that they are encoding claims and not from any modal fragility attaching to a contingent change in the expressive power of the language.

In any case, this is not a matter to be resolved here. Our system is prepared, whether or not we mark these imported analytic truths as modally fragile (and whether or not we tag any theorem derived from them with a $\star$).

15.6.2 Identifying Mathematical Individuals and Relations

(**983**) **Axiom:** The Reduction Axiom for Mathematical Individuals and Relations. The mathematical individuals and relations of a given theory may now be identified theoretically as follows. Where T is any mathematical theory, τ_T is any closed, well-defined term of type t added to the present theory in accordance with (981.2), $A!$ is a constant of type $\langle t\rangle$, and F is a variable to type $\langle t\rangle$, then it is axiomatic that τ_T is the abstract object of type t that encodes just the properties F such that, in theory T, τ_T exemplifies F:

(.1) **Reduction Axiom**: $\tau_T = \imath x(A!x \,\&\, \forall F(xF \equiv\ T \models F\tau_T))$

Here are instances of (.1) that identify mathematical individuals; they derive from the examples in (981.4), where the types on the expressions are obvious:

- $\varnothing_{\text{ZF}} = \imath x(A!x \,\&\, \forall F(xF \equiv\ \text{ZF} \models F\varnothing_{\text{ZF}}))$
- $\{\varnothing\}_{\text{ZF}} = \imath x(A!x \,\&\, \forall F(xF \equiv\ \text{ZF} \models F\{\varnothing\}_{\text{ZF}}))$
- $4_{\text{PA}} = \imath x(A!x \,\&\, \forall F(xF \equiv\ \text{PA} \models F4_{\text{PA}}))$
- $(2\times2)_{\text{PA}} = \imath x(A!x \,\&\, \forall F(xF \equiv\ \text{PA} \models F(2\times2)_{\text{PA}}))$
- $\pi_{\mathbb{R}} = \imath x(A!x \,\&\, \forall F(xF \equiv\ \mathbb{R} \models F\pi_{\mathbb{R}}))$

Here are some instances of (.1) that identify some of the primitive mathematical relations of type $\langle i,i\rangle$ and $\langle i\rangle$ involved in the examples in (981.4):

- $\in_{\text{ZF}} =\ \imath x(A!x \,\&\, \forall F(xF \equiv\ \text{ZF} \models F{\in_{\text{ZF}}}))$
- $S_{\text{ZF}} = \imath x(A!x \,\&\, \forall F(xF \equiv\ \text{ZF} \models FS_{\text{ZF}}))$
- $=_{\text{PA}}\ =\ \imath x(A!x \,\&\, \forall F(xF \equiv\ \text{PA} \models F{=_{\text{PA}}}))$
- $>_{\mathbb{R}}\ =\ \imath x(A!x \,\&\, \forall F(xF \equiv\ \mathbb{R} \models F{>_{\mathbb{R}}}))$

It should be clear that (a) the 1st, 3rd, and 4th relations identified above have type $\langle i, i\rangle$, (b) the 2nd has type $\langle i\rangle$, and (c) the types of all the other expressions can be inferred from these.

Finally, we produce instances of (.1) that identify some of the (complex) mathematical relations involved in the examples of (981.4):

- $[\lambda x\, x \in \{\varnothing\}]_{\mathrm{ZF}} \;=\; \imath x(A!x \,\&\, \forall F(xF \equiv \mathrm{ZF} \models F[\lambda x\, x \in \{\varnothing\}]_{\mathrm{ZF}}))$ type: $\langle i\rangle$
- $[\lambda F\, \varnothing F\{\varnothing\}]_{\mathrm{ZF}} = \imath x(A!x \,\&\, \forall G(xG \equiv \mathrm{ZF} \models G[\lambda F\, F\varnothing\{\varnothing\}]_{\mathrm{ZF}}))$ type: $\langle\langle i, i\rangle\rangle$
- $[\lambda yz\, \forall x(x > y \rightarrow x > z]_{\mathbb{R}} =$
  $\imath x(A!x \,\&\, \forall F(xF \equiv \mathbb{R} \models F[\lambda yz\, \forall x(x > y \rightarrow x > z]_{\mathbb{R}}))$ type: $\langle i, i\rangle$

It is important to recognize that these are *not* definitions of the objects in question but rather theoretical descriptions! The descriptions are well-defined because we've established that for each condition φ on properties with no free xs, there is a unique abstract object x that encodes just the properties satisfying φ. So the identity of the mathematical object in each case is ultimately secured by our ordinary mathematical judgements of the form $T \models Fx$, which themselves are grounded from facts of the form $\vdash_T Fx$.[413]

15.6.3 Theorems Governing Mathematical Entities

We now examine some consequences of (a) the new axioms that result from the Importation Principle (981), (b) the instances of the Reduction Axiom (983). We use these to derive general and particular facts about mathematical entities.

(**984**) **Theorems:** The Equivalence Principle and its Consequences. For any mathematical theory T and closed term τ_T of T of any type t, we may derive, as a general principle, that (.1) $\tau_T F$ if and only if (the proposition that) $F\tau_T$ is true in T:

(.1) **Equivalence Principle**: $\tau_T F \equiv T \models F\tau_T$

Clearly, the variable F in this principle will depend on the type of τ_T, and so will have type $\langle t\rangle$ whenever τ_T is a term of type t. It should be observed that since $T \models F\tau_T$ is an encoding claim, it is a necessary truth when true. So although this theorem is derivable by modally strict means, if it should turn out that the data imported in (981.4) are best understood as modally fragile

[413] The above identifications corrects the procedure in Zalta 2000b (SS4–6) and 2006b (676–7). In Zalta 2000b, we failed to index the λ-expressions; it wasn't yet clear that these should be interpreted as abstract properties and relations and so needed to be indexed to their respective theories. In Zalta 2006b, we mistakenly used the indexed terms $\{\varnothing_{\mathrm{ZF}}\}$ instead of $\{\varnothing\}_{\mathrm{ZF}}$. And on 677, we used the term $[\lambda x\, x \in \{\varnothing_{\mathrm{ZF}}\}]$ where we should have just used the term $[\lambda x\, x \in \{\varnothing\}]$. However, in 2006b, we weren't discussing the identity of mathematical relations, since the paper was focused on the notions of essence and modality as they applied to mathematical objects.

axioms, then any consequence of (.1) derived from such axioms will itself have to be flagged as a ⋆-theorem. For the remainder of this section, however, we may leave that question open, since the system is prepared for this eventuality.

The following encoding formulas are immediate consequences of (.1) and the data, i.e., the new axioms in (981); each formula is followed by a reading in *natural language with ambiguous predication*, which may be regarded as true when the copula 'has' is interpreted to mean 'encodes' rather than 'exemplifies'.

(.2) $\varnothing_{\text{ZF}}[\lambda x\, x \in \{\varnothing\}]_{\text{ZF}}$ [by (.1) and (981.4.c)]
The ZF emptyset has the ZF property: being an element of the unit set of the empty set.

(.3) $\{\varnothing\}_{\text{ZF}}[\lambda x\, \varnothing \in x]_{\text{ZF}}$ [by (.1) and (981.4.d)]
The ZF unit set of the empty set has the ZF property: having the empty set as a member.

(.4) $\in_{\text{ZF}}[\lambda F\, \varnothing F\{\varnothing\}]_{\text{ZF}}$ [by (.1) and (981.4.e)]
The ZF membership relation has the ZF property: being a relation that relates the empty set to the unit set of the empty set.

(.5) $\varnothing_{\text{ZF}}[\lambda y\, \neg\exists x(Sx \,\&\, x \in y)]_{\text{ZF}}$ [by (.1) and (981.4.h)]
The ZF empty set has the ZF property: having no set as a member.

(.6) $S_{\text{ZF}}[\lambda F\, \neg\exists x(Fx \,\&\, x \in \varnothing)]_{\text{ZF}}$ [by (.1) and (981.4.i)]
The ZF property of being a set has the ZF property: being a property such that nothing exemplifying it is an element of the empty set.

(.7) $\in_{\text{ZF}}[\lambda G\, \neg\exists x(Sx \,\&\, Gx\varnothing)]_{\text{ZF}}$ [by (.1) and (981.4.j)]
The ZF membership relation has the ZF property: being a relation that no set bears to the emptyset.

(.8) $=_{\text{PA}}[\lambda F\, \exists x(2{\times}2Fx)]_{\text{PA}}$ [F infix] [by (.1) and (981.4.l)]
The PA identity relation has the PA property: being a relation that relates 2×2 to something.

(.9) $\pi_{\mathbb{R}}[\lambda y\, \forall x(x > 4 \rightarrow x > y)]_{\mathbb{R}}$ [by (.1) and (981.4.n)]
The $\mathbb{R}$ number π has the $\mathbb{R}$ property: being an object than which everything greater than 4 is greater.

(.10) $>_{\mathbb{R}}[\lambda F\, \forall x(xF4 \rightarrow Fx\pi)]_{\mathbb{R}}$ [by (.1) and (981.4.p)]
The $\mathbb{R}$ greater-than relation has the $\mathbb{R}$ property: being a relation such that everything that bears it to 4 bears it to π.

And if we suppose that the sentence has, in each case, been uttered in a context that in which named mathematical theory is being discussed, then we can drop the theory-relative indices. Thus, (.2) becomes: the empty set has the property *being an element of the unit set of the empty set*; and (.3) becomes: the unit set of the empty set has the property *having the empty set as a member*; and so on. I take it that this shows that the theory preserves the data, while at the same time, makes it clear both as to what is denoted by the mathematical terms and as to the conditions under which the claim is true.

(**985**) **Remark:** How Math Objects Encode Math Relations. It seems reasonable to extend our systematization of n-ary encoding formulas so that we can state conditions under which mathematical objects encode mathematical relations. Axiom (935.29) asserts that an n-ary encoding formula $x_1 \ldots x_n F$ is true if and only and each of the x_i encodes $[\lambda y\, Fx_1 \ldots x_{i-1} y x_{i+1} \ldots x_n]$, for $1 \leq i \leq n$. Note that when we imported (981.4.c) and (981.4.d), the following became provable:

$$\varnothing_{\text{ZF}}[\lambda x\, x \in \{\varnothing\}]_{\text{ZF}} \tag{984.2}$$

$$\{\varnothing\}_{\text{ZF}}[\lambda x\, \varnothing \in x]_{\text{ZF}} \tag{984.3}$$

The conjunction of these two claims implies, by axiom (935.29), that $\varnothing_{\text{ZF}}$ and $\{\varnothing\}_{\text{ZF}}$ encode the relation $\in_{\text{ZF}}$, i.e., the following binary encoding formula of the form xyF is now a theorem:

$$\varnothing_{\text{ZF}}\{\varnothing\}_{\text{ZF}}\in_{\text{ZF}}$$

But notice also that we also imported (981.4.e), from which we proved:

$$\in_{\text{ZF}}[\lambda F\, \varnothing F\{\varnothing\}]_{\text{ZF}} \tag{984.4}$$

Then the conjunction of (984.2) – (984.4) implies, by axiom (935.29), the following ternary encoding formula of the form $xyzF$:

$$\varnothing_{\text{ZF}}\{\varnothing\}_{\text{ZF}}\in_{\text{ZF}}[\lambda xyF\, Fxy]_{\text{ZF}}$$

A similar move can be for more complex formulas, as follows.

When we imported (981.4.h) – (9814.j) and obtained (984.5) – (984.7), respectively, we then know the conjunction:

$$\varnothing_{\text{ZF}}[\lambda y\, \neg\exists x(Sx \,\&\, x \in y)]_{\text{ZF}} \;\&\; S_{\text{ZF}}[\lambda F\, \neg\exists x(Fx \,\&\, x \in \varnothing)]_{\text{ZF}} \;\&\; \in_{\text{ZF}}[\lambda G\, \neg\exists x(Sx \,\&\, Gx\varnothing)]_{\text{ZF}}$$

But then by axiom (935.29), it follows that:

$$\varnothing_{\text{ZF}}S_{\text{ZF}}\in_{\text{ZF}}[\lambda yFG\, \neg\exists x(Fx \,\&\, Gxy]_T$$

In general, when $\vdash_T \varphi$ and $\tau_1, \ldots, \tau_n$ are the well-defined, closed terms occurring in φ, we can prove, where the $z_1, \ldots, z_n$ are variables of the same type as $\tau_{1T}, \ldots, \tau_{nT}$, respectively:

$$\tau_{1T}[\lambda z_1\, \varphi^{z_1}_{\tau_{1T}}]_T \;\&\; \ldots \;\&\; \tau_{nT}[\lambda z_n\, \varphi^{z_n}_{\tau_{nT}}]_T$$

From this it follows by (935.29) that:

$$\tau_{1T} \ldots \tau_{nT}[\lambda z_1 \ldots z_n\, \varphi^{z_1,\ldots,z_n}_{\tau_{1T},\ldots,\tau_{nT}}]_T$$

Thus, we have a proof that abstract mathematical objects encode abstract mathematical relations.

(**986**) **Definition:** What Is It To Be an Object of a Mathematical Theory? Since we have taken a proof-theoretic approach to the analysis of mathematical language, we do not take the view that the objects of a theory T are those elements in the domain of the quantifiers of T. Rather, where x is a variable of any type t such that $t \neq \langle\,\rangle$, we can rest with the following definition: x is an *object of* T just in case it is true in T that x has a property. Formally, where we also use F as a variable of type $\langle t \rangle$:

(.1) $ObjectOf(x,T) \equiv_{df} \exists F(T \models Fx)$

When x has type i, we may call x an *individual* object of T, and when x has type $t \neq i$, i.e., when t is a type of the form $\langle t_1,\ldots,t_n\rangle$ $(n \geq 1)$, we may call x a *relational* object of T.

To see that (.1) yields the right results, consider the examples of the Importation Principle (981.4.c) and (981.4.e):

$$\mathrm{ZF} \models [\lambda x\, x \in \{\varnothing\}]_{\mathrm{ZF}} \varnothing_{\mathrm{ZF}}$$

$$\mathrm{ZF} \models [\lambda F\, \varnothing F\{\varnothing\}]_{\mathrm{ZF}} \in_{\mathrm{ZF}}$$

From these, it follows, respectively, that:

$$\exists F(\mathrm{ZF} \models F\varnothing_{\mathrm{ZF}})$$

$$\exists F(\mathrm{ZF} \models F\in_{\mathrm{ZF}})$$

From these last two axioms and (.1), we may infer both that $ObjectOf(\varnothing_{\mathrm{ZF}},\mathrm{ZF})$ and that $ObjectOf(\in_{\mathrm{ZF}},\mathrm{ZF})$, i.e., that $\varnothing_{\mathrm{ZF}}$ is an individual object of ZF and that $\in_{\mathrm{ZF}}$ is a relational object of ZF.

Moreover, (.1) correctly implies that two relations and no individuals are objects of the theory of dense, linear orderings without endpoints, which we discussed in (979) above. Let's call that theory D and import D into object theory, so that the theorems of D become axioms of the present theory. Then we know, for example:

$$D \models \forall x \forall y \forall z(x <_D y \,\&\, y <_D z \rightarrow x <_D z) \quad \text{(Transitivity)}$$
$$D \models \forall x \neg(x <_D x) \quad \text{(Irreflexivity)}$$
$$D \models \forall x \forall y(\neg(x =_D y) \rightarrow (x <_D y \vee y <_D x)) \quad \text{(Connectedness)}$$
$$D \models \forall x \forall y \exists z(x <_D z <_D y) \quad \text{(Dense)}$$
$$D \models \forall x \exists y \exists z(z <_D x <_D y) \quad \text{(No Endpoints)}$$

And we know that reflexivity and unrestricted substitution govern $=_D$, i.e.,

$D \models \forall x \forall y(x =_D y)$
$D \models \forall x \forall y(x =_D y \rightarrow (\varphi \rightarrow \varphi'))$, where φ' is the result of substituting zero or more occurrences of y for x in φ

So the only two well-defined, closed terms of D are the two primitive notions $<_D$ and $=_D$, both of type $\langle i, i\rangle$. It should also be clear that since both $<$ and $=$ are transitive in the theory D, the following are also the result of importing D:

$$D \models [\lambda F\, \forall x \forall y \forall z(xFy \,\&\, yFz \rightarrow xFz)]_D <_D$$

$$D \models [\lambda F\, \forall x \forall y \forall z(xFy \,\&\, yFz \rightarrow xFz)]_D =_D$$

Further, since the higher-order claims $\exists F(F<)$ and $\exists F(F=)$ are also theorems of D, we know:

$$D \models \exists F(F <_D)$$

$$D \models \exists F(F =_D)$$

So (.1) correctly predicts that $ObjectOf(<_D, D)$ and $ObjectOf(=_D, D)$. And since there are no closed and well-defined individual terms of D, we may conclude, by inspection, that there is no individual term τ for which we may conclude that $ObjectOf(\tau, D)$.

Though (.1) correctly represents the above facts, one should take care not to mistakenly conclude that non-unique witnesses to existential claims of T are objects of T. Nodelman & Zalta 2014 refined (.1) as follows, where x, y are variables of any type t ($t \neq \langle\,\rangle$) and F a variable of type $\langle t\rangle$:[414]

(.2) $ObjectOf(x, T) \equiv_{df} T \models \forall y(y \neq_T x \rightarrow \exists F \neg(Fx \equiv \neg Fy))$

This identifies the individual and relational objects of a theory as those which are are discernible in T. In that paper, we were concerned that someone might mistakenly attempt to argue that, given (.1), our analysis of complex number theory ($\mathbb{C}$) would imply $i_{\mathbb{C}} = -i_{\mathbb{C}}$. The mistaken argument would begin by observing that i and $-i$ are indiscernible in complex analysis, by a known automorphism of the complex plane. Thus, any open, i-free formula $\varphi(x)$ in $\mathbb{C}$ holds of i if and only if it holds of $-i$. So it would appear that the following accurately describes $\mathbb{C}$:

[414] The definition that follows uses a different definiendum than the one used in Nodelman & Zalta 2014. In that work, we identified a mathematical theory T with the structure T, since T encodes all and only the (quantified) propositions that indicate which objects stand in which relations. And so, instead of talking about the *objects* of the theory T, we talked instead about the *elements* of the structure T. But the idea is the same and so we reproduce here the definition of the elements of the structure T used in that paper, but now cast as a definition of the objects of the theory T.

(A) $\vdash_{\mathbb{C}} \forall F(Fi \equiv F{-}i)$

If so, then by the Importation Principle (981.3), it is a theorem of the present theory that:

(B) $\mathbb{C} \models \forall F(Fi_{\mathbb{C}} \equiv F{-}i_{\mathbb{C}})$

But is one were to assume that $i_{\mathbb{C}}$ and $-i_{\mathbb{C}}$ are objects of $\mathbb{C}$ and that this entitles one to identify them as abstract objects, one might then use the Reduction Axiom (983.1) and its consequence, the Equivalence Principle (984.1), to conclude:

(C) $\quad i_{\mathbb{C}}F \equiv \mathbb{C} \models Fi_{\mathbb{C}}$

(D) $-i_{\mathbb{C}}F = \mathbb{C} \models F{-}i_{\mathbb{C}}$

But (B), (C), and (D) would imply, in object theory, $i_{\mathbb{C}} = -i_{\mathbb{C}}$.[415] But, this seems to be in tension with the fact that $\mathbb{C} \models i_{\mathbb{C}} \neq_{\mathbb{C}} -i_{\mathbb{C}}$, which follows from the fact that $\vdash_{\mathbb{C}} i \neq -i$ by the Importation Principle (981.3).

The response in Nodelman & Zalta 2014 involved two steps, the first of which turns out to be unnecessary given the second. The first step was to (.2) instead of (.1) as the definition of the objects of a theory, so that for x to be an object of T, then for any other distinct object y of T, some property distinguishes x and y. The second step was to challenge (B) as a logically perspicuous (and ontologically relevant) theorem of $\mathbb{C}$, on the grounds that the expressions 'i' and '$-i$' in $\mathbb{C}$ aren't well-defined. We wrote:

> Indeed, we suggest that the correct procedure for interpreting the language of $\mathbb{C}$ is as follows: before importation, eliminate the logically non-well-defined term 'i' by replacing every theorem of the form $\varphi(\ldots i \ldots)$ by a theorem of the form: $\exists x(x^2 + 1 = 0 \,\&\, \varphi(\ldots x \ldots))$; then import the result. We suggest that this is the right procedure because mathematical practice here really involves two steps: (1) add the axiom that asserts $\exists x(x^2 + 1 = 0)$, and (2) eliminate the quantifier and introduce an arbitrary name for the existentially quantified variable. Though a structuralist should be happy enough with step (1), the use of arbitrary, non-well-defined names in step

[415] Consider an arbitrary property Q and assume $i_{\mathbb{C}}Q$. Then by (C):

(ϑ) $\mathbb{C} \models Qi_{\mathbb{C}}$

Now if we consider only derivability in $\mathbb{C}$ itself, then:

(ξ) $\forall F(Fi \equiv F{-}i),\, Qi \vdash_{\mathbb{C}} Q{-}i$

It follows from (ϑ), (ξ) and (B) by the Rule of Closure (981.6) that $\mathbb{C} \models Q{-}i_{\mathbb{C}}$. So by (D), $-i_{\mathbb{C}}Q$. Hence $i_{\mathbb{C}}Q \to -i_{\mathbb{C}}Q$. And by analogous reasoning, $-i_{\mathbb{C}}Q \to i_{\mathbb{C}}Q$. So $i_{\mathbb{C}}Q \equiv -i_{\mathbb{C}}Q$. Since Q was arbitrary, $\forall G(i_{\mathbb{C}}G \equiv -i_{\mathbb{C}}G)$. But since by hypothesis $i_{\mathbb{C}}$ and $-i_{\mathbb{C}}$ are abstract objects (this is a consequence of the Reduction Axiom), it would follow by the definition of abstract object identity that $i_{\mathbb{C}} = -i_{\mathbb{C}}$.

> (2) is not justified ontologically. Though we are quite happy to allow mathematical practice to carry on in the usual way, our view is that a philosopher may not appeal to that practice of using arbitrary names to generate *ontological* problems. Nodelman & Zalta 2014, 70–71

This procedure for handling i and $-i$ has now been subsumed by the normalization methods for formulating mathematical theories described in (979.5). These methods require us to replace axioms and theorems involving i or $-i$ in $\mathbb{C}$ with existentially quantified conjunctions of the kind described in the passage quoted above. Consequently, we don't add i_C and $-i_C$ to object theory as uniquely definable terms subject to the Reduction Axiom, though we may reason with them as long as they are considered *arbitrary* names to be discharged by some use of Existential Elimination (939) [102].

So given the methods in (979.5), we are not forced to replace definition (.1) with (.2). Our methods already ensure that i and $-i$ don't qualify as objects of $\mathbb{C}$. These methods don't yield the axioms $\mathbb{C} \models i_{\mathbb{C}} \neq_{\mathbb{C}} -i_{\mathbb{C}}$ or $\mathbb{C} \models \forall F(Fi_{\mathbb{C}} \equiv F{-i_{\mathbb{C}}})$. We've eliminated the uses of 'i' and '$-i_{\mathbb{C}}$' that might lead one to suppose these terms pose and ontological problem. So we can rest with (.1). We can't instantiate (.1) with the expressions '$i_{\mathbb{C}}$' or '$-i_{\mathbb{C}}$' to conclude that $ObjectOf(i_{\mathbb{C}},\mathbb{C})$ and $ObjectOf(-i_{\mathbb{C}},\mathbb{C})$, since we don't have $\vdash_{\mathbb{C}} \exists F(Fi)$ and $\vdash_{\mathbb{C}} \exists F(F{-i})$.

(**987**) **Remark:** A Note About Logical Completeness. Consider the Continuum Hypothesis (CH), where this is the claim $2^{\aleph_0} = \aleph_1$. Since it is logically true that $\mathrm{CH} \vee \neg\mathrm{CH}$, we know:

$$\vdash_{\mathrm{ZF}} (\mathrm{CH} \vee \neg\mathrm{CH})$$

This implies, by the Importation Principle, that OT is extended with the following analytic truth:

$$\mathrm{ZF} \models \big((2^{\aleph_0}{}_{\mathrm{ZF}} =_{\mathrm{ZF}} \aleph_{1\,\mathrm{ZF}}) \vee \neg(2^{\aleph_0}{}_{\mathrm{ZF}} =_{\mathrm{ZF}} \aleph_{1\,\mathrm{ZF}})\big)$$

which we may abbreviate, for simplicity, as:

$$\mathrm{ZF} \models (\mathrm{CH}_{\mathrm{ZF}} \vee \neg\mathrm{CH}_{\mathrm{ZF}})$$

However, *derivability* in ZF is not logically complete in the following sense: $\vdash_{\mathrm{ZF}} (\varphi \vee \psi)$ doesn't imply the disjunction: either $\vdash_{\mathrm{ZF}} \varphi$ or $\vdash_{\mathrm{ZF}} \psi$. CH is a case in point. Even though $\vdash_{\mathrm{ZF}} (\mathrm{CH} \vee \neg\mathrm{CH})$, the fact that CH is independent of the axioms of ZF just means that $\nvdash_{\mathrm{ZF}} \mathrm{CH}$ and $\nvdash_{\mathrm{ZF}} \neg\mathrm{CH}$. So it follows by metarule (981.5) that both $\neg\mathrm{ZF} \models \mathrm{CH}_{\mathrm{ZF}}$ and $\neg\mathrm{ZF} \models \neg\mathrm{CH}_{\mathrm{ZF}}$ are theorems of our theory. Thus, object theory has the following as a theorem:

$$\mathrm{ZF} \models (\mathrm{CH}_{\mathrm{ZF}} \vee \neg\mathrm{CH}_{\mathrm{ZF}})\ \&\ \neg(\mathrm{ZF} \models \mathrm{CH}_{\mathrm{ZF}})\ \&\ \neg(\mathrm{ZF} \models \neg\mathrm{CH}_{\mathrm{ZF}})$$

And generally, incomplete theories yield theorems of the form:

$$\text{ZF} \models (\varphi \vee \psi) \,\&\, \neg(\text{ZF} \models \varphi) \,\&\, \neg(\text{ZF} \models \psi)$$

The fact that a disjunction is true-in-T doesn't imply that one of the disjuncts is true-in-T. Finally, it should be apparent that the theorem $\neg\text{ZF} \models \text{CH}_{\text{ZF}}$ doesn't imply $\text{ZF} \models \neg\text{CH}_{\text{ZF}}$. This also offers a sense in which truth-in-T, is not logically complete.

(**988**) **Remark:** Two Remarks. (1) Why are we indexing only the primary terms when we import theorems of T into object theory? The reasons:

- If we index a term to T and then remove the index, then the result should be something in the language of T and not something in object theory. I.e., we should only index expressions that are expressions of T. Indexed terms are part of the language of object theory, not part of the language of T and if those indexed terms were to be subterms .
- Similarly, we don't want to index λ-expressions of object theory. If we indexed λ-expressions that contained indexed subterms, we would be indexing an expression of object theory, not of the target mathematical theory.
- We don't want to apply β-Conversion to any indexed λ-expression, since indexed λ-expression denote abstract relations. So this method removes the temptation to think β-Conversion applies to indexed λ-expressions.
- It is also easier to read and understand.

(2) Why aren't we analyzing 'data' containing open terms (e.g., open function terms or open descriptions, and open λ-expressions)? The answer: we may now consider it as legitimate, but with the understanding that no open term is to be indexed and analyzed. We can consider $\vdash_{\text{PA}} \exists z(z^2 = 4)$ as data, import this as $\text{PA} \models \exists z(z^2 =_{\text{PA}} 4_{\text{PA}})$. But we shall not attempt to give an 'open-ended' analysis such as:

$$z^2 = \imath x(A!x \,\&\, \forall F(xF \equiv \text{PA} \models Fz^2))$$

There is no data of the form $\text{PA} \models Fz^2$.

Part III

Metaphilosophy

Part IV

Technical Appendices, Bibliography, Index

Appendix: Proofs of Theorems and Metarules

NOTE: The items below numbered as (n), ($n.m$), or ($n.m.a$) refer to numbered items in Part II. So, for example, references to item (9.1) do *not* refer to Chapter 9.1, but rather to item (9.1), which occurs in Part II, Chapter 7.

(62.1) (Exercise)

(62.2) (Exercise)

(63.1) We establish the rule only for $\vdash$. If φ is an element of $\mathbf{\Lambda}$, then the one element sequence φ is a proof of φ, by (59.2). ⋈

(63.2) We establish the rule only for $\vdash$. If φ is an element of Γ, then the one element sequence φ is a derivation of φ from Γ, by (59.1). ⋈

(63.3) We establish the rule only for $\vdash$. If $\vdash \varphi$, then by definition (59.2), there is a sequence of formulas every element of which is either a member of $\mathbf{\Lambda}$ or a direct consequence of some of the preceding members of the sequence by virtue of MP. Since $\mathbf{\Lambda} \subseteq \mathbf{\Lambda} \cup \Gamma$, there is a sequence of formulas every element of which is either a member of $\mathbf{\Lambda} \cup \Gamma$ or a direct consequence from some of the preceding members of the sequence by virtue of MP. Hence, by definition (59.1), $\Gamma \vdash \varphi$. ⋈

(63.4) (Exercise)

(63.5) We establish the rule only for $\vdash$. Assume $\Gamma_1 \vdash \varphi$ and $\Gamma_2 \vdash (\varphi \to \psi)$. Then there is a sequence $\chi_1, \ldots, \chi_{n-1}, \varphi$ ($= S_1$) that is a derivation of φ from Γ_1 and there is a sequence $\theta_1, \ldots, \theta_{m-1}, \varphi \to \psi$ ($= S_2$) that is a derivation of $\varphi \to \psi$ from Γ_2. So the consider the sequence:

$$\chi_1, \ldots, \chi_{n-1}, \varphi, \theta_1, \ldots, \theta_{m-1}, \varphi \to \psi, \psi \qquad (S_3)$$

Since every element of S_1 and S_2 is either an element of $\mathbf{\Lambda} \cup \Gamma_1 \cup \Gamma_2$ or follows from preceding members by MP, the same holds for every member of the initial segment of S_3 up to and including $\varphi \to \psi$. Since the last member of S_3 follows from previous members by MP, we know that every element of S_3 is either an element of $\mathbf{\Lambda} \cup \Gamma_1 \cup \Gamma_2$ or follows from preceding members by MP. Hence, S_3 is a derivation of ψ from $\Gamma_1 \cup \Gamma 2$. This is a witness to $\Gamma_1, \Gamma_2 \vdash \psi$. ⋈

(63.6) (Exercise)

(63.7) We establish the rule only for $\vdash$. Assume $\Gamma \vdash \varphi$ and that $\Gamma \subseteq \Delta$. Then by the former, there is a sequence S ending in φ such that every member of the sequence is either in $\mathbf{\Lambda} \cup \Gamma$ or follows from previous members by MP. But since $\Gamma \subseteq \Delta$, it follows that every member of S is either in $\mathbf{\Lambda} \cup \Delta$ or follows from previous members by MP, i.e., $\Delta \vdash \varphi$. ⋈

(63.8) We establish the rule only for $\vdash$. Assume $\Gamma \vdash \varphi$ and $\varphi \vdash \psi$. By the former, there is a sequence $S_1 = \chi_1, \ldots, \chi_{n-1}, \varphi$ such that every member of the sequence is either in $\mathbf{\Lambda} \cup \Gamma$ or follows from previous members by MP. By the latter, there is a sequence $S_2 = \theta_1, \ldots, \theta_{m-1}, \psi$ such that every member of the sequence is either in $\mathbf{\Lambda} \cup \{\varphi\}$ or follows from previous members by MP. So consider, then, the following sequence:

$$\chi_1, \ldots, \chi_{n-1}, \theta_1, \ldots, \theta_{m-1}, \psi \tag{S_3}$$

This sequence is the concatenation of the first $n-1$ members of S_1 with the entire sequence S_2. Since S_1 and S_2 are derivations, we know that all the χ_i $(1 \leq i \leq n-1)$ and θ_j $(1 \leq j \leq m-1)$ are either elements of $\mathbf{\Lambda} \cup \Gamma$ or follow from two of the preceding members of the sequence by MP. The only potential exceptions are possible occurrences of φ among the θ_js. But note that φ follows by MP from two members of $\chi_1, \ldots, \chi_{n-1}$. Thus S_3 is a derivation of ψ from Γ and, hence, $\Gamma \vdash \psi$. ⋈

(63.9) We establish the rule only for $\vdash$. Suppose $\Gamma \vdash \varphi$. Since the instances of (38.1) are axioms, we know by (63.1) that $\vdash \varphi \to (\psi \to \varphi)$, where ψ is any formula. So by (63.3), we have $\Gamma \vdash \varphi \to (\psi \to \varphi)$. From our initial hypothesis and this last result, it follows by an instance of (63.5) (i.e., an instance in which we (a) set both Γ_1 and Γ_2 in (63.5) to Γ and (b) set ψ in (63.5) to $\psi \to \varphi$) that $\Gamma \vdash (\psi \to \varphi)$, where ψ is any formula. ⋈

(63.10) We establish the rule only for $\vdash$. Suppose $\Gamma \vdash (\varphi \to \psi)$. Since $\Gamma \subseteq \Gamma \cup \{\varphi\}$, it follows from (63.7) that:

(ϑ) $\Gamma \cup \{\varphi\} \vdash \varphi \to \psi$

But since $\varphi \in \Gamma \cup \{\varphi\}$, it follows by (63.2) that:

(ξ) $\Gamma \cup \{\varphi\} \vdash \varphi$

So from (ϑ) and (ξ), it follows by (63.5) (setting both Γ_1 and Γ_2 in (63.5) to Γ) that $\Gamma \cup \{\varphi\} \vdash \psi$, i.e., $\Gamma, \varphi \vdash \psi$. ⋈

(63.11) (Exercise)

(66) Suppose (a) $\Gamma \vdash \varphi$, and (b) α doesn't occur free in any formula in Γ. We show by induction on the length of the derivation of φ from Γ that $\Gamma \vdash \forall\alpha\varphi$.

Base case. The derivation of φ from Γ is a one-element sequence, in which case the sequence must be φ itself since a derivation of φ from Γ must end with φ. Then by the definition of *derivation from*, (59.1), $\varphi \in \mathbf{\Lambda} \cup \Gamma$. So we have two cases: (A) φ is an element of $\mathbf{\Lambda}$, i.e., φ is one of the axioms asserted in Chapter 8, or (B) φ is an element of Γ.

> *Case A*. $\varphi \in \mathbf{\Lambda}$. Then $\forall\alpha\varphi \in \mathbf{\Lambda}$, since we took the universal closures of all our axioms as axioms, i.e., the conditional, if $\varphi \in \mathbf{\Lambda}$, then $\forall\alpha\varphi \in \mathbf{\Lambda}$, governed our statement of the axioms. So, $\forall\alpha\varphi \in \mathbf{\Lambda} \cup \Gamma$, and so by (63.4), it follows that $\Gamma \vdash \forall\alpha\varphi$.
>
> *Case B*. $\varphi \in \Gamma$. Then, by hypothesis, α doesn't occur free in φ. Consequently, $\varphi \to \forall\alpha\varphi$ is an instance of axiom (39.4) meeting the condition that α doesn't occur free in φ. So by (59.1), the sequence $\varphi, \varphi \to \forall\alpha\varphi, \forall\alpha\varphi$ is a witness to $\Gamma \vdash \forall\alpha\varphi$, since every member of the sequence is either a member of $\mathbf{\Lambda} \cup \Gamma$ or is a direct consequence of two previous members by MP.

Inductive Case. Suppose that the derivation of φ from Γ is a sequence of length n, where $n > 1$. Then either $\varphi \in \mathbf{\Lambda} \cup \Gamma$ or φ follows from two previous members of the sequence, namely, $\psi \to \varphi$ and ψ, by MP. If $\varphi \in \mathbf{\Lambda} \cup \Gamma$, then using the same reasoning as in the base case, $\Gamma \vdash \forall\alpha\varphi$. If φ follows from previous members $\psi \to \varphi$ and ψ by MP, then by the definition of a derivation, we know that $\Gamma \vdash \psi \to \varphi$ and $\Gamma \vdash \psi$, where these are derivations of length less than n. Since our IH is that the theorem holds for all derivations of formulas from Γ of length less than n, it follows that $\Gamma \vdash \forall\alpha(\psi \to \varphi)$ and $\Gamma \vdash \forall\alpha\psi$. So there is a sequence $S_1 = \chi_1, \ldots, \chi_i$, where $\chi_i = \forall\alpha(\psi \to \varphi)$, that is a witness to the former and a sequence $S_2 = \theta_1, \ldots, \theta_j$, where $\theta_j = \forall\alpha\psi$, that is a witness to the latter. Now by using an instance of axiom (39.3), we may construct the following sequence:

$$\chi_1, \ldots, \chi_i, \theta_1, \ldots, \theta_j, \forall\alpha(\psi \to \varphi) \to (\forall\alpha\psi \to \forall\alpha\varphi), \forall\alpha\psi \to \forall\alpha\varphi, \forall\alpha\varphi \qquad (S_3)$$

The antepenultimate member of S_3 is an instance of axiom (39.3), and so an element of $\mathbf{\Lambda}$ and hence of $\mathbf{\Lambda} \cup \Gamma$. The penultimate member of S_3 follows from previous members (namely, the antepenultimate member and χ_i) by MP, and the last member of S_3 follows from previous members (namely, the penultimate member and θ_j) by MP. Hence, every element of S_3 is either in $\mathbf{\Lambda} \cup \Gamma$ or follows from previous members by MP. So $\Gamma \vdash \forall\alpha\varphi$. ⋈

(68) Suppose $\Gamma \vdash_{\Box} \varphi$, i.e., that there is a modally-strict derivation of φ from Γ. We show by induction on the length of the derivation that $\Box\Gamma \vdash_{\Box} \Box\varphi$, i.e., that there is a modally strict derivation of $\Box\varphi$ from $\Box\Gamma$.

Base Case. If $n = 1$, then the modally-strict derivation of φ from Γ consists of a single formula, namely, φ itself. So by the definition of $\Gamma \vdash_{\Box} \varphi$ (60), φ must be in $\mathbf{\Lambda}_{\Box} \cup \Gamma$. So we have two cases: (A) φ is in $\mathbf{\Lambda}_{\Box}$ or (B) φ is in Γ.

> *Case A*: $\varphi \in \mathbf{\Lambda}_{\Box}$. Then φ must be a necessary axiom and so its necessitation $\Box\varphi$ is an axiom. So $\vdash_{\Box} \Box\varphi$ by (63.1) and $\Box\Gamma \vdash_{\Box} \Box\varphi$ by (63.3).[426]
>
> *Case B*: $\varphi \in \Gamma$. Then $\Box\varphi$ is in $\Box\Gamma$, by the definition of $\Box\Gamma$ (68). Hence, by (63.2), it follows that $\Box\Gamma \vdash_{\Box} \Box\varphi$.

Inductive Case. Suppose that the modally-strict derivation of φ from Γ is a sequence S of length n, where $n > 1$. Then either $\varphi \in \mathbf{\Lambda}_{\Box} \cup \Gamma$ or φ follows by MP from two previous members of the sequence, namely, $\psi \to \varphi$ and ψ. If $\varphi \in \mathbf{\Lambda}_{\Box} \cup \Gamma$, then using the reasoning in the base case, it follows that $\Box\Gamma \vdash_{\Box} \Box\varphi$. If φ follows from previous members $\psi \to \varphi$ and ψ by MP, then by the definition of a modally-strict derivation, we know both that $\Gamma \vdash_{\Box} \psi \to \varphi$ and $\Gamma \vdash_{\Box} \psi$. Consequently, since our IH is that the theorem holds for all such derivations of length less than n, it implies:

(a) $\Box\Gamma \vdash_{\Box} \Box(\psi \to \varphi)$

(b) $\Box\Gamma \vdash_{\Box} \Box\psi$

Now since instances of the K schema (45.1) are necessary axioms (i.e., members of $\mathbf{\Lambda}_{\Box}$), we know by (63.1):

$$\vdash_{\Box} \Box(\psi \to \varphi) \to (\Box\psi \to \Box\varphi)$$

So by (63.3), it follows that:

$$\Box\Gamma \vdash_{\Box} \Box(\psi \to \varphi) \to (\Box\psi \to \Box\varphi)$$

So by (63.5) (setting both Γ_1 and Γ_2 in (63.5) to Γ), it follows from this and (a) that:

$$\Box\Gamma \vdash_{\Box} \Box\psi \to \Box\varphi$$

And again by (63.5), it follows from this and (b) that:

$$\Box\Gamma \vdash_{\Box} \Box\varphi$$ ⋈

(74) Axiom (38.2) asserts:

[426] Note that (63.3) says that if $\vdash_{\Box} \varphi$, then $\Gamma \vdash_{\Box} \varphi$, for *any* Γ. So, in this case, we've substituted $\Box\Gamma$ for Γ in (63.3). The clash of variables is not an egregious one.

$$\varphi \to (\psi \to \chi) \to ((\varphi \to \psi) \to (\varphi \to \chi))$$

If we let φ in the above be φ, let ψ in the above be $(\varphi \to \varphi)$, and let χ in the above be φ, then we obtain the following instance of (38.2):

$$(\varphi \to ((\varphi \to \varphi) \to \varphi)) \to ((\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi))$$

But the following is an instance of (38.1):

$$\varphi \to ((\varphi \to \varphi) \to \varphi)$$

Since this latter is the antecedent of the former, we may apply MP to obtain:

$$(\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi)$$

But now the following is also an instance of (38.1):

$$\varphi \to (\varphi \to \varphi)$$

By applying MP to our last two results we obtain:

$$\varphi \to \varphi \qquad \bowtie$$

(75) Suppose $\Gamma, \varphi \vdash \psi$. We show by induction on the length of a derivation of ψ from $\Gamma \cup \{\varphi\}$ that $\Gamma \vdash (\varphi \to \psi)$.

Base case. The derivation of ψ from $\Gamma \cup \{\varphi\}$ is a one-element sequence, namely, ψ itself. Then by the definition of *derivation from*, (59.1), $\psi \in \mathbf{\Lambda} \cup \Gamma \cup \{\varphi\}$. So we have two cases: (A) ψ is an element of $\mathbf{\Lambda} \cup \Gamma$, i.e., ψ is one of the axioms asserted in Chapter 8 or an element of Γ, or (B) $\psi = \varphi$.

Case A. $\psi \in \mathbf{\Lambda} \cup \Gamma$. Then by (59.1), $\Gamma \vdash \psi$. Since the instances of (38.1) are axioms governing conditionals, we know $\vdash (\psi \to (\varphi \to \psi))$, by (63.1). So, by (63.3), it follows that $\Gamma \vdash (\psi \to (\varphi \to \psi))$. Hence by (63.5), it follows that $\Gamma \vdash (\varphi \to \psi)$.

Case B. $\psi = \varphi$. Then by (74), we know $\vdash (\psi \to \psi)$. So, $\vdash (\varphi \to \psi)$, and hence, by (63.3), it follows that $\Gamma \vdash (\varphi \to \psi)$.

Inductive Case. The derivation of ψ from $\Gamma \cup \{\varphi\}$ is a sequence of length n, where $n > 1$. Then either $\psi \in \mathbf{\Lambda} \cup \Gamma \cup \{\varphi\}$ or ψ follows from two previous members of the sequence, namely, $\chi \to \psi$ and χ, by MP. If $\psi \in \mathbf{\Lambda} \cup \Gamma \cup \{\varphi\}$, then using the same reasoning as in the base case, $\Gamma \vdash (\varphi \to \psi)$. If ψ follows from previous members $\chi \to \psi$ and χ by MP, then since our IH is that the theorem holds for all derivations of formulas from Γ of length less than n, it implies both:

(a) $\Gamma \vdash (\varphi \to \chi)$

(b) $\Gamma \vdash (\varphi \to (\chi \to \psi))$

Now since the instances of (38.2) are axioms governing conditionals, we know, by (63.1):

$$\vdash (\varphi \to (\chi \to \psi)) \to ((\varphi \to \chi) \to (\varphi \to \psi))$$

So, by (63.3), it follows that:

$$\Gamma \vdash (\varphi \to (\chi \to \psi)) \to ((\varphi \to \chi) \to (\varphi \to \psi))$$

From this and (b), it follows by (63.5) that:

$$\Gamma \vdash (\varphi \to \chi) \to (\varphi \to \psi)$$

And from this last conclusion and (a), it follows that:

$$\Gamma \vdash (\varphi \to \psi) \qquad \bowtie$$

(76.1) Assume:

(a) $\Gamma_1 \vdash \varphi \to \psi$

(b) $\Gamma_2 \vdash \psi \to \chi$

So, by definition (59.1), there is a sequence, say S_1, that is a witness to (a) and a sequence, say S_2, that is a witness to (b). Then consider the sequence S_3 consisting of the members of S_1, followed by the members of S_2, followed by $\varphi \to \psi$, $\psi \to \chi$, φ, ψ, and ending in χ. It is not hard to show that this is a witness to:

(ϑ) $\Gamma_1, \Gamma_2, \varphi \to \psi, \psi \to \chi, \varphi \vdash \chi$

since every element of S_3 either: (a) is an element of Γ_1, or (b) is an element of Γ_2, or (c) is just the formula $\varphi \to \psi$, $\psi \to \chi$, or φ, or (d) follows from previous members of the sequence by MP. By an application of the Deduction Theorem to (ϑ), it follows that:

$$\Gamma_1, \Gamma_2, \varphi \to \psi, \psi \to \chi \vdash \varphi \to \chi$$

By an application of the Deduction Theorem to the above, and another application to the result, we obtain:

(ξ) $\Gamma_1, \Gamma_2, \vdash (\varphi \to \psi) \to ((\psi \to \chi) \to (\varphi \to \chi))$

But from (a) and (b), respectively, it follows by (63.7) that:

(c) $\Gamma_1, \Gamma_2 \vdash \varphi \to \psi$

(d) $\Gamma_1, \Gamma_2 \vdash \psi \to \chi$

So from (ξ) and (c) it follows by (63.5) that:

(ζ) $\Gamma_1, \Gamma_2, \vdash (\psi \to \chi) \to (\varphi \to \chi)$

And from (ζ) and (d) it again follows by (63.5) that $\Gamma_1, \Gamma_2 \vdash \varphi \to \chi$. ⋈

(76.2) (Exercise)

(76.3) Consider the premise set $\Gamma = \{\varphi \to \psi,\ \psi \to \chi,\ \varphi\}$. From the first and third members of Γ, we obtain ψ by MP. From ψ and the second member of Γ, we obtain χ by MP. Hence, the sequence consisting of the members of Γ followed by ψ and χ constitute a witness to $\varphi \to \psi,\ \psi \to \chi,\ \varphi \vdash \chi$. So by the Deduction Theorem (75), $\varphi \to \psi,\ \psi \to \chi \vdash \varphi \to \chi$. ⋈

(76.4) Consider the premise set $\Gamma = \{\varphi \to (\psi \to \chi),\ \psi,\ \varphi\}$. Then from the first and third members of Γ, we obtain $\psi \to \chi$ by MP, and from this and the second member of Γ we obtain χ by MP. Hence the sequence consisting of the members of Γ followed by $\psi \to \chi$ and χ constitute a witness to $\varphi \to (\psi \to \chi),\ \psi,\ \varphi \vdash \chi$. So by the Deduction Theorem (75), it follows that $\varphi \to (\psi \to \chi),\ \psi \vdash \varphi \to \chi$. ⋈

(77.1) As an instance of (38.3), we have: $(\neg\varphi \to \neg\neg\varphi) \to ((\neg\varphi \to \neg\varphi) \to \varphi)$. Moreover, by (74), we have $\neg\varphi \to \neg\varphi$. Then by (76.4), it follows that $(\neg\varphi \to \neg\neg\varphi) \to \varphi$. But $(\neg\neg\varphi \to (\neg\varphi \to \neg\neg\varphi))$ is an instance of (38.1). So it follows that $\neg\neg\varphi \to \varphi$, by (76.1). ⋈

(77.2) As an instance of (38.3), we have: $(\neg\neg\neg\varphi \to \neg\varphi) \to ((\neg\neg\neg\varphi \to \varphi) \to \neg\neg\varphi)$. Moreover, as an instance of (77.1), we know: $\neg\neg\neg\varphi \to \neg\varphi$. So by MP, it follows that $(\neg\neg\neg\varphi \to \varphi) \to \neg\neg\varphi$. But as an instance of (38.1), we know: $\varphi \to (\neg\neg\neg\varphi \to \varphi)$. So by (76.1), it follows that $\varphi \to \neg\neg\varphi$. ⋈

(77.3) Assume $\neg\varphi$ for conditional proof. Now assume φ for a conditional proof nested within our conditional proof. Then since $\varphi \to (\neg\psi \to \varphi)$ is an instance of axiom (38.1), it follows from this and our second assumption that:

(a) $\neg\psi \to \varphi$

Moroever, since $\neg\varphi \to (\neg\psi \to \neg\varphi)$ is an instance of axiom (38.1), it follows from this and our first assumption that:

(b) $\neg\psi \to \neg\varphi$

But as an instance of axiom (38.3), we know:

(c) $(\neg\psi \to \neg\varphi) \to ((\neg\psi \to \varphi) \to \psi)$

From (b) and (c), it follows that $(\neg\psi \to \varphi) \to \psi$. And from this and (a), it follows that ψ. So, discharging the premise of our nested conditional proof, it follows that $\varphi \to \psi$. Hence, discharging the premise of our original conditional proof, it follows that $\neg\varphi \to (\varphi \to \psi)$. ⋈

(77.4) We establish $(\neg\psi \to \neg\varphi) \to (\varphi \to \psi)$ by conditional proof. Assume $\neg\psi \to \neg\varphi$. Then as an instance of (38.3), we know: $(\neg\psi \to \neg\varphi) \to ((\neg\psi \to \varphi) \to \psi)$. So it follows that $(\neg\psi \to \varphi) \to \psi$. But as an instance of (38.1), we know: $\varphi \to (\neg\psi \to \varphi)$. So by hypothetical syllogism (76.1) from from our last two results, it follows that $\varphi \to \psi$. So, by conditional proof (CP), it follows that $(\neg\psi \to \neg\varphi) \to (\varphi \to \psi)$. ⋈

(77.5) We establish $(\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$ by conditional proof. Assume $\varphi \to \psi$. We know by (77.1) that $\neg\neg\varphi \to \varphi$. So it follows by hypothetical syllogism (76.3) that:

(a) $\neg\neg\varphi \to \psi$

But by (77.2), we know:

(b) $\psi \to \neg\neg\psi$

So it follows from (a) and (b) by hypothetical syllogism (76.3) that $\neg\neg\varphi \to \neg\neg\psi$. But as an instance of (77.4), we know: $(\neg\neg\varphi \to \neg\neg\psi) \to (\neg\psi \to \neg\varphi)$. Hence it follows that $\neg\psi \to \neg\varphi$. So, by conditional proof (CP), it follows that $(\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$. ⋈

(77.6) Assume $\varphi \to \neg\psi$, to show $\psi \to \neg\varphi$ by conditional proof. Now assume ψ for a conditional proof nested within our conditional proof. From ψ it follows by (77.1) that $\neg\neg\psi$. Then from $\varphi \to \neg\psi$ and $\neg\neg\psi$, it follows by Modus Tollens that $\neg\varphi$. So discharging the premise of our nested conditional proof, we have $\psi \to \neg\varphi$. And discharging the premise of our original conditional proof, it follows that $(\varphi \to \neg\psi) \to (\psi \to \neg\varphi)$. ⋈

(77.7) (Exercise)

(77.8) – **(77.9)** Follow the proofs in Mendelson 1964 [1997, 39–40, Lemma 1.11(f) – (g)]. ⋈

(77.10) (Exercise)

(78.1) – **(78.2)** (Exercises)

(79.1) Assume $\Gamma_1 \vdash (\varphi \to \psi)$ and $\Gamma_2 \vdash \neg\psi$. Since $\Gamma_1 \subseteq \Gamma_1 \cup \Gamma_2$, it follows from the first assumption by (63.7) that:

(a) $\Gamma_1, \Gamma_2 \vdash (\varphi \to \psi)$

Since $\Gamma_2 \subseteq \Gamma_1 \cup \Gamma_2$, it follows from the second assumption by (63.7) that:

(b) $\Gamma_1, \Gamma_2 \vdash \neg\psi$

Now as an instance of (77.5), we know:

$$\vdash (\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$$

and hence by (63.3) that:

$$\Gamma_1, \Gamma_2 \vdash (\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$$

So by applying (63.5) to this last result and (a), we have: $\Gamma_1, \Gamma_2 \vdash \neg\psi \to \neg\varphi$. And by applying (63.5) to this result and (b), we have $\Gamma_1, \Gamma_2 \vdash \neg\varphi$. ⋈

(79.2) (Exercise)

(80.1) (→) Assume:

(ϑ) $\Gamma \vdash \varphi \to \psi$

But given (77.5), we know:

$$\vdash (\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$$

and hence by (63.3) that:

$$\Gamma \vdash (\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$$

Hence by applying (63.5) to this last result and (ϑ), we obtain $\Gamma \vdash (\neg\psi \to \neg\varphi)$. (←) By symmetric reasoning, but using (77.4). ⋈

(80.2) (→) Assume:

(ϑ) $\Gamma \vdash \varphi \to \neg\psi$

But given (77.6), we know:

$$\vdash (\varphi \to \neg\psi) \to (\psi \to \neg\varphi)$$

and hence by (63.3) that:

$$\Gamma \vdash (\varphi \to \neg\psi) \to (\psi \to \neg\varphi)$$

Hence by applying (63.5) to this last result and (ϑ), we obtain $\Gamma \vdash (\psi \to \neg\varphi)$. (←) By symmetric reasoning, but using (77.7) ⋈

(81.1) Assume $\Gamma_1, \neg\varphi \vdash \neg\psi$ and $\Gamma_2, \neg\varphi \vdash \psi$. By analogy with the first step of the reasoning in (79.1), it follows by (63.7) that both:

(a) $\Gamma_1, \Gamma_2, \neg\varphi \vdash \neg\psi$

(b) $\Gamma_1, \Gamma_2, \neg\varphi \vdash \psi$

Now, by the Deduction Theorem (75), it follows from (a) and (b), respectively, that:

(ϑ) $\Gamma_1, \Gamma_2 \vdash (\neg\varphi \to \neg\psi)$

(ζ) $\Gamma_1, \Gamma_2 \vdash (\neg\varphi \to \psi)$

But the instances of (38.3) are axioms and hence theorems, by (63.1). So we know:

$$\vdash (\neg\varphi \to \neg\psi) \to ((\neg\varphi \to \psi) \to \varphi)$$

From this result it follows by (63.3) that:

(ξ) $\Gamma_1, \Gamma_2 \vdash (\neg\varphi \to \neg\psi) \to ((\neg\varphi \to \psi) \to \varphi)$

But by apply (63.5) to (ϑ) and (ξ); to the result and (ζ), apply (63.5) again. It follows that $\Gamma_1, \Gamma_2 \vdash \varphi$. ⋈

(81.2) (Exercise)

(83) As an instance of (74), we know $\neg\varphi \to \neg\varphi$. But by the definition of $\vee$ (18.2) and the Rule of Definition by Equivalence (72), we also know that the following is a theorem:

$$(\neg\varphi \to \neg\varphi) \to (\varphi \vee \neg\varphi)$$

Hence, by Modus Ponens, $\varphi \vee \neg\varphi$. ⋈

(84) By the definition of & (18.1) and the Rule of Definition by Equivalence (72), we know that the following is a theorem:

$$(\varphi \,\&\, \neg\varphi) \to \neg(\varphi \to \neg\neg\varphi) \qquad ⋈$$

If we label $\varphi \,\&\, \neg\varphi$ as A, and label $\varphi \to \neg\neg\varphi$ as B, then the theorem displayed above has the form $A \to \neg B$. Now if we can show $A \to B$, then by a form of Reductio Ad Absurdum, namely, the Variant of (81.2), it follows that $\neg A$, which is what we want to prove. But we already know B, by theorem (77.2), which asserts $\varphi \to \neg\neg\varphi$. And axiom (38.1) tells us if something is already established, then any claim whatsoever implies it. Hence $A \to B$. ⋈

(85.1) For conditional proof, assume $\varphi \,\&\, \psi$, to show φ. Now by definition of & (18.1) and the Rule of Definition by Equivalence (72), we know that $(\varphi \,\&\, \psi) \to \neg(\varphi \to \neg\psi)$ is a theorem. From this and our assumption, it follows that $\neg(\varphi \to \neg\psi)$. Hence, by axiom (38.1), everything implies $\neg(\varphi \to \neg\psi)$ and so, in particular, we may conclude:

(ϑ) $\neg\varphi \to \neg(\varphi \to \neg\psi)$

Independently, assume $\neg\varphi$, for a nested conditional proof. Then by (77.3), it follows that φ implies any formula, and so $\varphi \to \neg\psi$. Hence, by our nested conditional proof:

(ξ) $\neg\varphi \to (\varphi \to \neg\psi)$

From (ϑ) and (ξ) we may infer φ, by a form of Reductio Ad Absurdum, namely, the Variant of (81.1). So, by conditional proof, $(\varphi \,\&\, \psi) \to \varphi$. ⋈

(85.2) For conditional proof, assume $\varphi \,\&\, \psi$, to show ψ. Now by definition of & (18.1) and the Rule of Definition by Equivalence (72), we know that $(\varphi \,\&\, \psi) \to \neg(\varphi \to \neg\psi)$ is a theorem. From this and our assumption, it follows that $\neg(\varphi \to \neg\psi)$. Hence, by axiom (38.1), everything implies $\neg(\varphi \to \neg\psi)$ and so, in particular, we may conclude:

(ϑ) $\neg\psi \to \neg(\varphi \to \neg\psi)$

Independently, assume $\neg\psi$, for a nested conditional proof. Then by axiom (38.1), it follows that $\varphi \to \neg\psi$. Hence, by our nested conditional proof:

(ξ) $\neg\psi \to (\varphi \to \neg\psi)$

From (ϑ) and (ξ) we may infer ψ, by a form of Reductio Ad Absurdum, namely, the Variant of (81.1). So, by conditional proof, $(\varphi \,\&\, \psi) \to \psi$. ⋈

(85.3) By (77.2), we know:

(A) $\varphi \to \neg\neg\varphi$

Independently, by (77.3), we know:

(B) $\neg\neg\varphi \to (\neg\varphi \to \psi)$

So by hypothetical syllogism (76.3) from (A) and (B), it follows that:

(C) $\varphi \to (\neg\varphi \to \psi)$

Independently, by definition of $\vee$ (18.2) and the Rule of Definition by Equivalence (72), we also know the following is a theorem:

(D) $(\neg\varphi \to \psi) \to (\varphi \vee \psi)$

So by hypothetical syllogism from (C) and (D) it follows that $\varphi \to (\varphi \vee \psi)$. ⋈

(85.4) (Exercise)

(85.5) Assume φ, for conditional proof, to show $\psi \to (\varphi \,\&\, \psi)$. Assume ψ, for conditional proof, to show $\varphi \,\&\, \psi$. By definition of & (18.1) and the Rule of Definition by Equivalence (72), we know that $\neg(\varphi \to \neg\psi) \to (\varphi \,\&\, \psi)$ is a theorem. So, to show $\varphi \,\&\, \psi$, it suffices by Modus Ponens to show $\neg(\varphi \to \neg\psi)$. We prove the latter by appealing to the Variant version of Reductio Ad Absurdum (81.2) and, in particular, by showing both:

(ϑ) $(\varphi \to \neg\psi) \to \neg\psi$

(ξ) $(\varphi \to \neg\psi) \to \psi$

We establish (ϑ) by conditional proof. Suppose $\varphi \to \neg\psi$. Then from this and φ (our first assumption), it follows that $\neg\psi$. We establish (ξ) immediately from our second assumption, ψ, by axiom (38.1). Hence, by reductio, $\neg(\varphi \to \neg\psi)$, which sufficed to show $\varphi \,\&\, \psi$. And so by conditional proof, $\psi \to (\varphi \,\&\, \psi)$, and by conditional proof, $\varphi \to (\psi \to (\varphi \,\&\, \psi))$. ⋈

(85.6) Let A be $(\varphi \,\&\, \varphi)$ and B be φ. Then since we want to prove $A \equiv B$, we may use the following proof strategy:

1. Prove $A \to B$
2. Prove $B \to A$
3. Independently, we know that theorem (85.5) has an instance of the form:

   $$(A \to B) \to ((B \to A) \to ((A \to B) \,\&\, (B \to A)))$$

4. From steps 1, 2 and 3, it follows that:

   $$(A \to B) \,\&\, (B \to A)$$

5. By definition of $\equiv$ (18.3) and the Rule of Definition by Equivalence (72), we know the following is a theorem:

   $$((A \to B) \,\&\, (B \to A)) \to (A \equiv B)$$

6. From steps 4 and 5 it follows that $A \equiv B$

Given this strategy, it remains only to show steps 1 and 2. Step 1 is easy, since $(\varphi \,\&\, \varphi) \to \varphi$ is an instance of (85.1). We prove step 2 by conditional proof. So assume φ. Then as an instance of (85.5), we know $\varphi \to (\varphi \to (\varphi \,\&\, \varphi))$. But then by two applications of Modus Ponens, it follows that $\varphi \,\&\, \varphi$. ⋈

(85.7) Let A be $(\varphi \vee \varphi)$ and B be φ. Then since we want to prove $A \equiv B$, we may use the proof strategy we used in (85.6). So it remains only to show steps 1 and 2.

We prove step 1 by conditional proof. So assume $\varphi \vee \varphi$. Now by definition of $\vee$ (18.2) and the Rule of Definition by Equivalence (72), we know $(\varphi \vee \varphi) \to (\neg\varphi \to \varphi)$ is a theorem. Hence $\neg\varphi \to \varphi$. But independently we know, as an instance of theorem (74), that $\neg\varphi \to \neg\varphi$. So it follows from our last two results that φ, by Reductio Ad Absurdum (81.1).

Step 2 is easy, since it is an instance of (85.3). ⋈

(86.1) Assume $\Gamma_1 \vdash \varphi$ and $\Gamma_2 \vdash \psi$. Note independently, by theorem (85.5), that $\varphi \to (\psi \to (\varphi \,\&\, \psi))$. Hence, $\Gamma_1 \vdash \varphi \to (\psi \to (\varphi \,\&\, \psi))$, by (63.3). So from our first assumption and this last result, it follows by (63.5) that $\Gamma_1 \vdash \psi \to (\varphi \,\&\, \psi)$. But from this and our second assumption, it follows that $\Gamma_1, \Gamma_2 \vdash \varphi \,\&\, \psi$, by (63.5).

⋈ [NOTE: In the usual manner, the justification of the $\vdash_{\Box}$ form of the rule is analogous.]

(86.2.a) – **(86.3.b)** (Exercises)

(86.3.c) Assume:

(A) $\Gamma_1 \vdash \varphi \lor \psi$

(B) $\Gamma_2 \vdash \varphi \to \chi$

(C) $\Gamma_3 \vdash \psi \to \theta$

We want to show $\Gamma_1, \Gamma_2, \Gamma_3 \vdash \chi \lor \theta$. Note that by definition of $\lor$ (18.2) and the Rule of Definition by Equivalence (72), we know $\vdash (\varphi \lor \psi) \to (\neg\varphi \to \psi)$ is a theorem. Hence by (63.3):

(D) $\Gamma_1 \vdash (\varphi \lor \psi) \to (\neg\varphi \to \psi)$

So from (A) and (D) it follows by (63.5) that:

(E) $\Gamma_1 \vdash \neg\varphi \to \psi$

Now independently, it follows from (B) by a rule of contraposition (80.1) that:

(F) $\Gamma_2 \vdash \neg\chi \to \neg\varphi$

But then from (F) and (E) it follows by a corollary to the Deduction Theorem, namely (76.1), that:

(G) $\Gamma_2, \Gamma_1 \vdash \neg\chi \to \psi$

But from (G) and (C) it follows by the same corollary that:

(H) $\Gamma_2, \Gamma_1, \Gamma_3 \vdash \neg\chi \to \theta$

But we know, by definition of $\lor$ (18.2) and the Rule of Definition by Equivalence (72), that $\vdash (\neg\chi \to \theta) \to (\chi \lor \theta)$ is a theorem. So by (63.3):

(I) $\Gamma_2, \Gamma_1, \Gamma_3 \vdash (\neg\chi \to \theta) \to (\chi \lor \theta)$

So from (H) and (I) it follows by (63.5) that $\Gamma_2, \Gamma_1, \Gamma_3 \vdash \chi \lor \theta$. But given (63.11), the order in which premise sets are listed makes no difference, and so $\Gamma_1, \Gamma_2, \Gamma_3 \vdash \chi \lor \theta$. ⋈

(86.4.a) Assume $\Gamma_1 \vdash \varphi \lor \psi$, $\Gamma_2 \vdash \varphi \to \chi$, and $\Gamma_3 \vdash \psi \to \chi$. We want to show $\Gamma_1, \Gamma_2, \Gamma_3 \vdash \chi$. Note that if we let θ in (86.3.c) be χ, then we know:

If $\Gamma_1 \vdash \varphi \lor \psi$, $\Gamma_2 \vdash \varphi \to \chi$, and $\Gamma_3 \vdash \psi \to \chi$, then $\Gamma_1, \Gamma_2, \Gamma_3 \vdash \chi \lor \chi$.

Hence it follows that:

(ϑ) $\Gamma_1, \Gamma_2, \Gamma_3 \vdash \chi \vee \chi$

Independently, by the idempotency of $\vee$ (85.7), it is a theorem that $(\chi \vee \chi) \equiv \chi$. By definition of $\equiv$ (18.3) and the Rule of Definition by Equivalence (72), this implies $((\chi \vee \chi) \rightarrow \chi) \,\&\, (\chi \rightarrow (\chi \vee \chi))$. So by Rule &E (86.2), it follows that $(\chi \vee \chi) \rightarrow \chi$. Since this is a theorem, we know $\vdash (\chi \vee \chi) \rightarrow \chi$, and so by (63.3), it follows that:

(ξ) $\Gamma_1, \Gamma_2, \Gamma_3 \vdash (\chi \vee \chi) \rightarrow \chi$

So, from (ξ) and (ϑ), it follows that $\Gamma_1, \Gamma_2, \Gamma_3 \vdash \chi$, by (63.5). ⋈

(86.4.b) – **(86.4.c)** (Exercises)

(87.1) Assume $\Gamma, \neg\varphi \vdash \psi \,\&\, \neg\psi$. But $\neg(\psi \,\&\, \neg\psi)$ is a theorem, i.e., $\vdash \neg(\psi \,\&\, \neg\psi)$, by (84). Hence, by (63.3), $\neg(\psi \,\&\, \neg\psi)$ follows from *any* premise set. So, in particular: $\Gamma, \neg\varphi \vdash \neg(\psi \,\&\, \neg\psi)$. But then, by our original form of Reductio Ad Absurdum (81.1), it follows that $\Gamma \vdash \varphi$. ⋈

(87.2) – **(87.6)** (Exercises)

(88.1) – **(88.2.f)** (Exercises)

(88.3.a) By the definition of $\equiv$ (18.3) and the Rule of Definition by Equivalence (72), it suffices to show: $(\varphi \rightarrow \varphi) \,\&\, (\varphi \rightarrow \varphi)$. But $\varphi \rightarrow \varphi$ is a theorem (74), and so by the idempotence of & (85.6), it follows that $(\varphi \rightarrow \varphi) \,\&\, (\varphi \rightarrow \varphi)$. ⋈

(88.3.b) (Exercise)

(88.3.c) We use reductio (87.2) to prove $\neg(\varphi \equiv \neg\varphi)$, by assuming $\varphi \equiv \neg\varphi$ and deriving a contradiction of the form $\chi \,\&\, \neg\chi$. It follows from our assumption, by the definition of $\equiv$ (18.3) and the Rule of Definition by Equivalence (72), that:

$$(\varphi \rightarrow \neg\varphi) \,\&\, (\neg\varphi \rightarrow \varphi)$$

So by &E, we know both:

(ϑ) $\varphi \rightarrow \neg\varphi$

(ξ) $\neg\varphi \rightarrow \varphi$

Note that $(\varphi \rightarrow \neg\varphi) \rightarrow ((\varphi \rightarrow \varphi) \rightarrow \neg\varphi)$ is an instance of (77.10). From this instance and (ϑ), it follows that $(\varphi \rightarrow \varphi) \rightarrow \neg\varphi$. But $\varphi \rightarrow \varphi$ is a theorem (74). Hence $\neg\varphi$. But from this last result and (ξ) it follows that φ. By Rule &I, we may conclude $\neg\varphi \& \varphi$ and, by the commutativity of & (88.2.a), conclude $\varphi \& \neg\varphi$, which is a contradiction. ⋈

(88.4.a) – **(88.8.h)** (Exercises)

(88.8.i) Assume $\varphi \equiv (\psi \,\&\, \chi)$, for conditional proof, to show $\psi \rightarrow (\varphi \equiv \chi)$. By definition of $\equiv$ and the Rule of Definition by Equivalence (72), our assumption implies:

$$(\varphi \rightarrow (\psi \,\&\, \chi)) \,\&\, ((\psi \,\&\, \chi) \rightarrow \varphi)$$

So by Rules &E (86.2.a) and (86.2.b), this last result implies:

(ϑ) $\varphi \rightarrow (\psi \,\&\, \chi)$

(ξ) $(\psi \,\&\, \chi) \rightarrow \varphi$

We want to show $\psi \rightarrow (\varphi \equiv \chi)$. So assume ψ, to show $\varphi \equiv \chi$. By definition of $\equiv$ and our Rule of Definition by Equivalence, it suffices to show $(\varphi \rightarrow \chi) \,\&\, (\chi \rightarrow \varphi)$. By Rule &I (86.1), it suffices to show each conjunct.

For the first conjunct, assume φ, for conditional proof. Then from φ and (ϑ), it follows that $\psi \,\&\, \chi$. Hence, by Rule &E (86.2.b), χ.

For the second conjunct, assume χ, for conditional proof. From our assumption ψ and our assumption χ, it follows by Rule &I that $\psi \,\&\, \chi$. Hence, by (ξ), it follows that φ. ⋈

(89.1) – **(89.3.f)** (Exercises)

(90.1) [Note: As usual, we justify only the $\vdash$ form of the rule; the proof of the $\vdash_\Box$ form is analogous.] Suppose $\varphi \equiv_{df} \psi$ is an instance of a definition-by-$\equiv$. Then by the Rule of Definition by Equivalence (72), we know both $\vdash \varphi \rightarrow \psi$ and $\vdash \psi \rightarrow \varphi$. So by (63.3), $\Gamma \vdash \varphi \rightarrow \psi$ and $\Gamma \vdash \psi \rightarrow \varphi$. But then, by Rule &I (86.1), it follows that:

(ϑ) $\Gamma \vdash (\varphi \rightarrow \psi) \,\&\, (\psi \rightarrow \varphi)$

Independently, by the definition of $\equiv$ and the Rule of Definition by Equivalence, we know that $\vdash ((\varphi \rightarrow \psi) \,\&\, (\psi \rightarrow \varphi)) \rightarrow (\varphi \equiv \psi)$. Hence, by (63.3):

(ξ) $\Gamma \vdash ((\varphi \rightarrow \psi) \,\&\, (\psi \rightarrow \varphi)) \rightarrow (\varphi \equiv \psi)$

Hence, from (ϑ) and (ξ), it follows by (63.5) that $\Gamma \vdash \varphi \equiv \psi$. ⋈

(90.2) Let $\varphi \equiv_{df} \psi$ be an instance of a definition-by-$\equiv$. Then by the Rule of Equivalence by Definition (90.1) we know $\Gamma \vdash \varphi \equiv \psi$. Now assume $\Gamma \vdash \varphi$. Hence by biconditional syllogism, i.e., Rule $\equiv$E (89.3.a), it follows that $\Gamma \vdash \psi$. ⋈

(90.3) (Exercise)

(91.1) Assume $\Gamma \vdash \varphi \equiv (\psi \,\&\, \chi)$ and $\Gamma \vdash \psi$, to show $\Gamma \vdash \varphi \equiv \chi$. Then by Rule &I (86.1), it follows that:

(ϑ) $\Gamma \vdash (\varphi \equiv (\psi \,\&\, \chi)) \,\&\, \psi$

Independently, it follows from (88.8.i) and the relevant instance of (88.7.b) that:

$$((\varphi \equiv (\psi \,\&\, \chi)) \,\&\, \psi) \rightarrow (\varphi \equiv \chi)$$

From this it follows by (63.3) that:

(ξ) $\Gamma \vdash ((\varphi \equiv (\psi \,\&\, \chi)) \,\&\, \psi) \rightarrow (\varphi \equiv \chi)$

From (ϑ) and (ξ) it follows that $\Gamma \vdash \varphi \equiv \chi$. ⋈

(91.2) (Exercise)

(93.1) Assume $\Gamma_1 \vdash \forall\alpha\varphi$ and $\Gamma_2 \vdash \tau\!\downarrow$. Assume further that τ is substitutable for α in φ. By (63.7), it follows both that:

(a) $\Gamma_1, \Gamma_2 \vdash \forall\alpha\varphi$

(b) $\Gamma_1, \Gamma_2 \vdash \tau\!\downarrow$

Independently, since we've taken instances of (39.1) as axioms, we know by (63.1) that:

$$\vdash \forall\alpha\varphi \rightarrow (\tau\!\downarrow \rightarrow \varphi^{\tau}_{\alpha})$$

Hence by (63.3), we know:

$$\Gamma_1, \Gamma_2 \vdash \forall\alpha\varphi \rightarrow (\tau\!\downarrow \rightarrow \varphi^{\tau}_{\alpha})$$

From this, (a) and (63.5), it follows that:

$$\Gamma_1, \Gamma_2 \vdash \tau\!\downarrow \rightarrow \varphi^{\tau}_{\alpha}$$

But from this, (b) and (63.5), it follows that $\Gamma_1, \Gamma_2 \vdash \varphi^{\tau}_{\alpha}$. ⋈

(93.1) [Proof of the Variant form of the rule.] Assume $\forall\alpha\varphi$ and $\tau\!\downarrow$, where τ is substitutable for α in φ. Then by (86.1), it follows that:

$$\forall\alpha\varphi \,\&\, \tau\!\downarrow$$

However, by applying an appropriate instance of Exportation (88.7.b) to our first quantifier axiom (39.1), we know that:

$$(\forall\alpha\varphi \,\&\, \tau\!\downarrow) \rightarrow \varphi^{\tau}_{\alpha}$$

So by MP, φ^{τ}_{α}. Thus, we've established $\forall\alpha\varphi, \tau\!\downarrow \vdash \varphi^{\tau}_{\alpha}$. ⋈

(93.2) Assume that $\Gamma \vdash \forall\alpha\varphi$, that τ is substitutable for α in φ, and that τ is either a primitive constant, a variable, or a core λ-expression. Then $\tau\!\downarrow$ is an axiom, by (39.2). So by (63.1), $\vdash \tau\!\downarrow$, and hence, $\Gamma \vdash \tau\!\downarrow$ by (63.3). So by (93.1) (with $\Gamma = \Gamma_1 = \Gamma_2$), it follows that $\Gamma \vdash \varphi^{\tau}_{\alpha}$. ⋈

(93.2) [Proof of the Variant form of the rule.] Assume $\forall\alpha\varphi$ and suppose both that τ is substitutable for α in φ and that τ is a primitive constant, a variable, or a core λ-expression. Then, by (39.2), we know $\tau\downarrow$. Once we conjoin $\forall\alpha\varphi$ and $\tau\downarrow$ by &I, it follows by reasoning used in the proof of the preceding theorem that φ^{τ}_{α}. Hence we've established that $\forall\alpha\varphi \vdash \varphi^{\tau}_{\alpha}$. ⋈

(95.1) Assume $\forall\alpha\varphi$, that τ is substitutable for α in φ, and that τ is either a primitive constant, a variable, or a core λ-expression. Then by Rule $\forall$E (93.2), it follows that φ^{τ}_{α}. So by conditional proof (CP), $\forall\alpha\varphi \to \varphi^{\tau}_{\alpha}$. ⋈

(95.2) Assume $\forall\alpha(\varphi \to \psi)$, where α is not free in φ. Now assume φ for a nested conditional proof. Since α isn't free in φ, it follows from φ by axiom (39.4) that $\forall\alpha\varphi$. So from our initial assumption $\forall\alpha(\varphi \to \psi)$ and this last result, it follows by axiom (39.3) that $\forall\alpha\psi$. Hence we may discharge the premise of our nested conditional proof to conclude $\varphi \to \forall\alpha\psi$, and then discharge the premise of our initial conditional proof to conclude that: $\forall\alpha(\varphi \to \psi) \to (\varphi \to \forall\alpha\psi)$.[427] ⋈

(95.3) (Exercise)

(96) Follow the proof of Theorem 24F in Enderton 1972 [2001, 123–124]. All of the results needed to complete the proof are in place. [Warning: Note that in the proof of this theorem, Enderton uses φ^{c}_{y} to be the result of replacing every occurrence of the constant c in φ with an occurrence of the variable y. By contrast, we use φ^{α}_{τ} to be the result of replacing every occurrence of the constant τ in φ with an occurrence of the variable α.] ⋈

(97.1) We need not prove this lemma by induction; the recursive definitions of φ^{τ}_{α} and *substitutable for* carry the following reasoning through all the inductive cases. Assume β is substitutable for α in φ and β doesn't occur free in φ. There are two cases. In the case where φ has no free occurrences of α, then by the definition of *substitutable for*, β is trivially substitutable for α in φ. In that case, however, φ^{β}_{α} just is φ and so $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta} = \varphi^{\alpha}_{\beta}$. Moreover by hypothesis, β doesn't occur free in φ. So by an analogous fact, $\varphi^{\alpha}_{\beta} = \varphi$. Hence $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta} = \varphi$.

In the case where α has at least one free occurrence in φ, then without loss of generality, consider any free occurrence of α in φ. Since β is substitutable for α in φ (by hypothesis), we know that β will not be bound when substituted for α at this occurrence. Thus, β will be free at this occurrence in φ^{β}_{α}. And since φ has no free occurrences of β (by hypothesis), we know that every free occurrence of β in φ^{β}_{α} replaced a free occurrence of α in φ. Thus no free occurrence of β in φ^{β}_{α} falls under the scope of a variable binding operator that binds α. Hence, α is substitutable for β in φ^{β}_{α}.

Now we must show that $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta} = \varphi$. Suppose, for reductio, that $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta} \neq \varphi$. Since substitution only changes the substituted variables, then there must be

[427] I'm indebted to Wes Anderson for noticing an error in a previous version of this proof.

some occurrence of α in $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$ that is not in φ or some occurrence of α in φ that is not in $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$. But both cases lead to contradiction:

> Case 1. Suppose there is an occurrence of α in $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$ that is not in φ. Then there was an occurrence of β in φ that remained an occurrence of β in φ^{β}_{α} but was replaced by an occurrence of α in $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$. But if an occurrence of β in φ^{β}_{α} was replaced by an occurrence of α in $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$, then that occurrence of β in φ^{β}_{α} had to be a free occurrence. But that occurrence of β must have been a free occurrence in φ (had it been a bound occurrence, it would have remained a bound occurrence in φ^{β}_{α}). And this contradicts the hypothesis that β doesn't occur free in φ.
>
> Case 2. Suppose there is an occurrence of α in φ that is not in $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$. Then there was an occurrence of α in φ that was replaced by an occurrence of β in φ^{β}_{α} which, in turn, remained an occurrence of β in $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$. But if that occurrence of β remained an occurrence of β in the re-replacement, then it must be bound by a variable-binding operator binding β in φ^{β}_{α}. But this contradicts the hypothesis that β is substitutable for α in φ, which requires that β must remain free at every occurrence of α in φ that it replaces in φ^{β}_{α}. ⋈

(97.2) Assume τ is a constant symbol that doesn't occur in φ. If α has no free occurrences in φ, then τ is trivially substitutable for α in φ and $\varphi^{\tau}_{\alpha} = \varphi$. In that case, both $(\varphi^{\tau}_{\alpha})^{\beta}_{\tau} = \varphi$, and $\varphi^{\beta}_{\alpha} = \varphi$. Hence $(\varphi^{\tau}_{\alpha})^{\beta}_{\tau} = \varphi^{\beta}_{\alpha}$. So we consider only the case where α has at least one free occurrence in φ.

Suppose, for reductio, that $(\varphi^{\tau}_{\alpha})^{\beta}_{\tau} \neq \varphi^{\beta}_{\alpha}$. Then there must be some occurrence of β in $(\varphi^{\tau}_{\alpha})^{\beta}_{\tau}$ that is not in φ^{β}_{α} or some occurrence of β in φ^{β}_{α} that is not in $(\varphi^{\tau}_{\alpha})^{\beta}_{\tau}$. But both cases lead to contradiction:

> Case 1. Suppose there is an occurrence of β in $(\varphi^{\tau}_{\alpha})^{\beta}_{\tau}$ that is not in φ^{β}_{α}. Since by hypothesis, there are no occurrences of τ in φ, then there must be an occurrence of α in φ that remained an occurrence of α in φ^{τ}_{α} but became an occurrence of β in $(\varphi^{\tau}_{\alpha})^{\beta}_{\tau}$. But if an occurrence of α in φ^{τ}_{α} became an occurrence of β in $(\varphi^{\tau}_{\alpha})^{\beta}_{\tau}$, then that occurrence of α in φ^{τ}_{α} had to be a free occurrence. But that contradicts the fact that φ^{τ}_{α} is, by definition, the result of replacing every free occurrence of α by τ in φ.
>
> Case 2. Suppose there is an occurrence of β in φ^{β}_{α} that is not in $(\varphi^{\tau}_{\alpha})^{\beta}_{\tau}$. Then there must be an occurrence of α in φ that became an occurrence of τ in φ^{τ}_{α} which remained an occurrence of τ in $(\varphi^{\tau}_{\alpha})^{\beta}_{\tau}$. But that contradicts the definition of ψ^{β}_{τ}, which signifies the result of replacing every occurrence of τ in ψ by an occurrence of β. ⋈

(97.3) (Exercise)

(99.1) $(\rightarrow)$ Assume $\forall\alpha\forall\beta\varphi$, for conditional proof (to show $\forall\beta\forall\alpha\varphi$). Then by the special case [Variant] of Rule $\forall$E (93.3), it follows that $\forall\beta\varphi$. By a second application of this same rule it follows that φ. Since α is not free in our hypothesis $\forall\alpha\forall\beta\varphi$, we may infer $\forall\alpha\varphi$ from φ, by GEN. Since β is not free in our hypothesis, we may infer $\forall\beta\forall\alpha\varphi$ from $\forall\alpha\varphi$, by GEN. Hence, by conditional proof, $\forall\alpha\forall\beta\varphi \rightarrow \forall\beta\forall\alpha\varphi$. $(\leftarrow)$ By symmetric reasoning. ⋈

(99.2) (Exercise)

(99.3) Assume $\forall\alpha(\varphi \equiv \psi)$ and apply the special case [Variant] of Rule $\forall$E (93.3) to obtain $\varphi \equiv \psi$. By $\equiv$I (89.2), it suffices to establish both directions of $\forall\alpha\varphi \equiv \forall\alpha\psi$. $(\rightarrow)$ Assume $\forall\alpha\varphi$. So by the special case [Variant] of Rule $\forall$E (93.3), we have φ. By a biconditional syllogism (89.3.a), it follows that ψ. Since α is not free in either of our premises, it follows that $\forall\alpha\psi$, by GEN. Discharging our second assumption, we've established $\forall\alpha\varphi \rightarrow \forall\alpha\psi$. $(\leftarrow)$ Assume $\forall\alpha\psi$. The conclusion is then reached by analogous reasoning, but by biconditional syllogism (89.3.b). ⋈

(99.4) $(\rightarrow)$ Assume, for conditional proof, that $\forall\alpha(\varphi \,\&\, \psi)$. Then by the special case [Variant] of Rule $\forall$E (93.3), we have $\varphi \,\&\, \psi$. From this we have both φ and ψ, by (86.2.a) and (86.2.b), respectively. Since α isn't free in our assumption, we may apply GEN to both conclusions to obtain $\forall\alpha\varphi$ and $\forall\alpha\psi$. Hence by (86.1), it follows that $\forall\alpha\varphi \,\&\, \forall\alpha\psi$. [We here omit the last step of assembling the conditional to be proved, since it is now obvious.] $(\leftarrow)$ Assume $\forall\alpha\varphi \,\&\, \forall\alpha\psi$, for conditional proof. It follows by (86.2.a) and (86.2.b) that $\forall\alpha\varphi$ and $\forall\alpha\psi$. Hence, by applying the special case [Variant] of Rule $\forall$E (93.3) to both, we obtain both φ and ψ. So by &I, we have $\varphi \,\&\, \psi$. Since α isn't free in our assumption, we may apply GEN to obtain $\forall\alpha(\varphi \,\&\, \psi)$. ⋈

(99.5) Assume, for conditional proof, that $\forall\alpha_1 \ldots \forall\alpha_n\varphi$. By the special case [Variant] of Rule $\forall$E (93.3), it follows that $\forall\alpha_2 \ldots \forall\alpha_n\varphi$. By analogous reasoning, we can strip off the quantifier $\forall\alpha_2$. Once we have legitimately stripped off the outermost quantifier in this way a total of n times, it follows that φ. ⋈

(99.6) $(\rightarrow)$ This direction is an instance of theorem (95.3). $(\leftarrow)$ Assume $\forall\alpha\varphi$. Then since α isn't free in our assumption, we may apply GEN to obtain $\forall\alpha\forall\alpha\varphi$. So by conditional proof, $\forall\alpha\varphi \rightarrow \forall\alpha\forall\alpha\varphi$. ⋈

(99.7) By hypothesis, α isn't free in φ. By $\equiv$I (89.2), it suffices to prove both directions of the biconditional. $(\rightarrow)$ Assume $\varphi \rightarrow \forall\alpha\psi$. Now for a secondary conditional proof, assume φ. Then by MP, it follows that $\forall\alpha\psi$ and by the special [Variant] case of Rule $\forall$E (93.3), it follows that ψ. Discharging the premise of our secondary conditional proof, it follows that $\varphi \rightarrow \psi$. Since α isn't free in φ, it isn't free in our remaining (original) assumption. So the conditions of GEN are met and we may conclude that $\forall\alpha(\varphi \rightarrow \psi)$. $(\leftarrow)$ By (95.2). ⋈

(99.8) (Exercise)

(99.9) (Exercise)

(99.10) Assume $\forall\alpha(\varphi \equiv \psi) \,\&\, \forall\alpha(\psi \equiv \chi)$. By &E, this yields $\forall\alpha(\varphi \equiv \psi)$ and $\forall\alpha(\psi \equiv \chi)$. By the special case [Variant] of Rule $\forall$E (93.3), it follows, respectively, that $\varphi \equiv \psi$ and $\psi \equiv \chi$. By a biconditional syllogism (89.3.e), it follows that $\varphi \equiv \chi$. Since α isn't free in our assumption, we may apply GEN to conclude $\forall\alpha(\varphi \equiv \chi)$. ⋈

(99.11) (Exercise)

(99.12) (Exercise)

(99.13) Suppose β is substitutable for α in φ and doesn't occur free in φ. Note that β could still be α, since (a) α is trivially substitutable for α in any φ, and (b) α may not occur free in φ. So we have two cases:

Case 1. β just is the variable α. Then our theorem becomes: $\forall\alpha\varphi \equiv \forall\alpha\varphi^{\alpha}_{\alpha}$. But $\varphi^{\alpha}_{\alpha}$ just is φ, by definition. So our theorem becomes $\forall\alpha\varphi \equiv \forall\alpha\varphi$, which is an instance of a tautology.

Case 2. β is distinct from α. ($\rightarrow$) Assume $\forall\alpha\varphi$. Since β is a variable and substitutable for α in φ, it follows by Rule $\forall$E (93.2) that φ^{β}_{α}. Furthermore, since β doesn't occur free in φ, β doesn't occur free in our assumption $\forall\alpha\varphi$. So we may apply GEN to obtain $\forall\beta(\varphi^{\beta}_{\alpha})$. ($\leftarrow$) Assume $\forall\beta(\varphi^{\beta}_{\alpha})$. Since β is substitutable for α in φ and doesn't occur free in φ, it follows, by the re-replacement lemma (97.1), both (a) that α is substitutable for β in φ^{β}_{α} and (b) that $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta} = \varphi$. From (a) and the fact that α is a variable, it follows from our assumption that $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$, by Rule $\forall$E (93.2). But by (b), this is just φ, and since α isn't free in our assumption, it follows by GEN that $\forall\alpha\varphi$. ⋈

(99.14) Let $\alpha_1,\ldots,\alpha_n$ be any distinct variables, for $n \geq 2$. As an instance of axiom (39.3), we know:

$$\forall\alpha_n(\varphi \rightarrow \psi) \rightarrow (\forall\alpha_n\varphi \rightarrow \forall\alpha_n\psi)$$

Since this is a theorem, it follows by GEN that:

$$\forall\alpha_{n-1}(\forall\alpha_n(\varphi \rightarrow \psi) \rightarrow (\forall\alpha_n\varphi \rightarrow \forall\alpha_n\psi))$$

By axiom (39.3), we may distribute $\forall\alpha_{n-1}$ over the conditional to conclude:

$$(\vartheta)\quad \forall\alpha_{n-1}\forall\alpha_n(\varphi \rightarrow \psi) \rightarrow \forall\alpha_{n-1}(\forall\alpha_n\varphi \rightarrow \forall\alpha_n\psi)$$

But as an instance of (39.3), we know:

$$(\xi)\quad \forall\alpha_{n-1}(\forall\alpha_n\varphi \rightarrow \forall\alpha_n\psi) \rightarrow (\forall\alpha_{n-1}\forall\alpha_n\varphi \rightarrow \forall\alpha_{n-1}\forall\alpha_n\psi)$$

So by biconditional syllogism from (ϑ) and (ξ), it follows that:

$$\forall\alpha_{n-1}\forall\alpha_n(\varphi \to \psi) \to (\forall\alpha_{n-1}\forall\alpha_n\varphi \to \forall\alpha_{n-1}\forall\alpha_n\psi)$$

If $n = 2$, we're done; otherwise if we apply this same series of reasoning steps to the variables $\alpha_{n-2},\ldots,\alpha_1$, we obtain:

$$\forall\alpha_1\ldots\forall\alpha_n(\varphi \to \psi) \to (\forall\alpha_1\ldots\forall\alpha_n\varphi \to \forall\alpha_1\ldots\forall\alpha_n\psi) \qquad \bowtie$$

(99.15) (Exercise)

(100) Assume $\Gamma \vdash \varphi^\tau_\alpha$, where τ is a primitive constant (i.e., isn't introduced by a definition) of the same type as α that doesn't occur in Γ, $\mathbf{\Lambda}$, or φ. Then by Rule $\forall$I (96), it follows that:

$$\Gamma \vdash \forall\beta(\varphi^\tau_\alpha)^\beta_\tau$$

where β is a variable (of the same type as α and τ) that doesn't occur in φ^τ_α. Since τ, by hypothesis, doesn't occur in Γ or φ, we know by the Re-replacement Lemma (97.2) that:

$$(\varphi^\tau_\alpha)^\beta_\tau = \varphi^\beta_\alpha$$

Thus, from our last two displayed results, it follows that:

$$\Gamma \vdash \forall\beta\varphi^\beta_\alpha$$

But independently, it follows from the right-to-left direction of (99.13) that:

$$\forall\beta\varphi^\beta_\alpha \vdash \forall\alpha\varphi$$

by (63.10). Hence, from our last two displayed results, it follows that:

$$\Gamma \vdash \forall\alpha\varphi \qquad \bowtie$$

(101.1) For simplicity, we justify the Variant version of the rule. By hypothesis, τ is substitutable for α in φ. Our proof strategy is:

(A) Show $(\varphi^\tau_\alpha \;\&\; \tau{\downarrow}) \to \exists\alpha\varphi$

(B) Conclude $\varphi^\tau_\alpha \to (\tau{\downarrow} \to \exists\alpha\varphi)$ by Exportation (88.7.a)

(C) Apply (63.10) twice to (B) to conclude: $\varphi^\tau_\alpha, \tau{\downarrow} \vdash \exists\alpha\varphi$

So it remains to show (A). Assume both φ^τ_α and $\tau{\downarrow}$. Note that the following is an instance of axiom (39.1):

$$\forall\alpha\neg\varphi \to (\tau{\downarrow} \to \neg\varphi^\tau_\alpha)$$

We leave it as an exercise to show that this is equivalent to:

$$\tau{\downarrow} \to (\forall\alpha\neg\varphi \to \neg\varphi^\tau_\alpha)$$

It follows from this theorem and our second assumption that $\forall\alpha\neg\varphi \rightarrow \neg\varphi^{\tau}_{\alpha}$. Hence, by contraposition (80.2), $\varphi^{\tau}_{\alpha} \rightarrow \neg\forall\alpha\neg\varphi$. From this and our first assumption, it follows that $\neg\forall\alpha\neg\varphi$. But then by the Rule of Definiendum Introduction, i.e., Rule $\equiv_{df}$I (90.3), it follows from this last conclusion and definition (18.4) that $\exists\alpha\varphi$. ⋈

(101.2) (Exercise)

(102) Assume $\Gamma, \varphi^{\tau}_{\alpha} \vdash \psi$, where τ is a primitive constant that does not occur in φ, ψ, Γ, or $\mathbf{\Lambda}$. Then we leave it as an exercise to show the 'contrapositive', i.e., that:

$$\Gamma, \neg\psi \vdash \neg\varphi^{\tau}_{\alpha}$$

where $\neg\varphi^{\tau}_{\alpha}$ is short for $\neg(\varphi^{\tau}_{\alpha})$. Since $\neg(\varphi^{\tau}_{\alpha}) = (\neg\varphi)^{\tau}_{\alpha}$ by metadefinition (14), we henceforth ignore the distinct ways of describing the same formula. Now since τ is, by hypothesis, a primitive constant that does not occur in φ, ψ, Γ, or $\mathbf{\Lambda}$, we know independently by the Variant of the Corollary to Rule $\forall$I (100), that:

$$\neg\varphi^{\tau}_{\alpha} \vdash \forall\alpha\neg\varphi$$

So by (63.8), it follows from our two displayed results that $\Gamma, \neg\psi \vdash \forall\alpha\neg\varphi$. So by the Deduction Theorem (75):

$$\Gamma \vdash \neg\psi \rightarrow \forall\alpha\neg\varphi$$

It follows, by the Rule of Contraposition (80):

$$\Gamma \vdash \neg\forall\alpha\neg\varphi \rightarrow \neg\neg\psi$$

Now given theorem (77.1), i.e., $\neg\neg\psi \rightarrow \psi$, and (63.3), we independently know:

$$\Gamma \vdash \neg\neg\psi \rightarrow \psi$$

So by Hypothetical Syllogism (76.1), our last two displayed results imply:

$$\Gamma \vdash \neg\forall\alpha\neg\varphi \rightarrow \psi$$

Now, independently, by definition of $\exists$ (18.4) and our Rule of Definition by Equivalence (72), we know:

$$\Gamma \vdash \exists\alpha\varphi \rightarrow \neg\forall\alpha\neg\varphi$$

So by Hypothetical Syllogism, our last two displayed results imply:

$$\Gamma \vdash \exists\alpha\varphi \rightarrow \psi$$

Hence, by (63.10), $\Gamma, \exists\alpha\varphi \vdash \psi$. ⋈

(103.1) Assume $\forall\alpha\varphi$. Then by the special case [Variant] of Rule $\forall$E (93.3), it follows that φ. Since φ has the form $\varphi^{\alpha}_{\alpha}$ and α is both a variable and substitutable for itself in φ, the conditions of Rule $\exists$I (101.2) are satisfied and we may infer $\exists\alpha\varphi$. ⋈

(103.2) By $\equiv$I (89.2), it suffices to prove both directions of the biconditional. ($\rightarrow$) Assume $\neg\forall\alpha\varphi$. We want to show $\exists\alpha\neg\varphi$. So by the definition of $\exists$ (18.4) and our Rule of Definiendum Introduction (90.3), we have to show: $\neg\forall\alpha\neg\neg\varphi$. For reductio, assume $\forall\alpha\neg\neg\varphi$. Then it follows that $\neg\neg\varphi$, by the special case [Variant] of Rule $\forall$E (93.3). Hence, by double negation elimination (78.2), we may infer φ. Since α isn't free in any of our assumptions, it follows by GEN that $\forall\alpha\varphi$. Since we've reached a contradiction, we may discharge our reductio assumption and conclude, by a version of RAA (87.4), that $\neg\forall\alpha\neg\neg\varphi$, which is all that remained to show. ($\leftarrow$) [For this direction of the proof, we shall appeal to Rule $\exists$E (102) in the form in which it is stated in the text. However, in subsequent proofs, we shall often use Rule $\exists$E in the manner described in the paragraph immediately following the introduction of the rule.] Assume, where τ is arbitrary (i.e., some fresh, primitive constant) that $\neg\varphi^{\tau}_{\alpha}$. Now assume, for reductio, that $\forall\alpha\varphi$. Since τ is a primitive constant, it is substitutable for α in φ. So by Rule $\forall$E (93.2), we have φ^{τ}_{α}. Contradiction. So by RAA (87.4), $\neg\forall\alpha\varphi$. We've thus shown $\neg\varphi^{\tau}_{\alpha} \vdash \neg\forall\alpha\varphi$. Since τ doesn't appear in φ or $\neg\forall\alpha\varphi$ or in any of our axioms, we may apply Rule $\exists$E (102) to conclude $\exists\alpha\neg\varphi \vdash \neg\forall\alpha\varphi$. ⋈

(103.3) By $\equiv$I (89.2), it suffices to prove both directions of the biconditional. [Henceforth, we omit mention of this proof strategy for biconditionals.] ($\rightarrow$) Assume $\forall\alpha\varphi$. We want to show $\neg\exists\alpha\neg\varphi$. For reductio, assume $\exists\alpha\neg\varphi$. From this and (103.2), it follows that $\neg\forall\alpha\varphi$, by a biconditional syllogism (89.3.b). Contradiction. So by RAA (87.3), $\neg\exists\alpha\neg\varphi$. ($\leftarrow$) Assume $\neg\exists\alpha\neg\varphi$, for conditional proof. Assume, for reductio, that $\neg\forall\alpha\varphi$. From this and (103.2), it follows that $\exists\alpha\neg\varphi$, by a biconditional syllogism (89.3.a). Contradiction. So by RAA (87.4), $\forall\alpha\varphi$. ⋈

(103.4) (Exercise)

(103.5) (Exercise)

(103.6) (Exercise)

(103.7) (Exercise)

(103.8) (Exercise)

(103.9) Assume $\neg\exists\alpha\varphi \;\&\; \neg\exists\alpha\psi$. From the first conjunct and (103.4) it follows that $\forall\alpha\neg\varphi$, and from the second conjunct and the same theorem it follows that $\forall\alpha\neg\psi$. Hence, by respective applications of the special case [Variant] of Rule

$\forall$E (93.3), it follows that $\neg\varphi$ and $\neg\psi$, which by &I gives us $\neg\varphi \,\&\, \neg\psi$. By $\vee$I (86.3.b), we may infer $(\varphi \,\&\, \psi) \vee (\neg\varphi \,\&\, \neg\psi)$. So by (88.4.g), we know $\varphi \equiv \psi$. Since α isn't free in our assumption, it follows by GEN that $\forall\alpha(\varphi \equiv \psi)$. $\bowtie$

(103.10) (Exercise)

(103.11) (Exercise)

(104.1) The definition of proposition existence (20.3) is:

$$p{\downarrow} \equiv_{df} [\lambda x\, p]{\downarrow}$$

By Conventions (17.2.a) and (17.2.b), all the variables in this definition function as metavariables. So the definition could be rewritten as:

$$\Pi^0{\downarrow} \equiv_{df} [\lambda \nu\, \Pi^0]{\downarrow}, \text{ provided } \nu \text{ isn't free in } \Pi^0$$

So let Π^0 be any 0-ary relation term, and choose ν to be some individual variable not free in Π^0. Then, by the Rule of Equivalence by Definition (90), it follows from the above definition that:

$$\Pi^0{\downarrow} \equiv [\lambda \nu\, \Pi^0]{\downarrow}$$

But since ν isn't free in Π^0, $[\lambda \nu\, \Pi^0]$ is a core λ-expression and so axiom (39.2) asserts $[\lambda \nu\, \Pi^0]{\downarrow}$. Hence, $\Pi^0{\downarrow}$. $\bowtie$

(104.2) By the previous theorem (104.1), $\Pi^0{\downarrow}$, for any relation term Π^0. But by the BNF definition in (4), all and only formulas φ are 0-ary relation terms (4). So $\varphi{\downarrow}$. $\bowtie$

(106) By (39.2), the modal closures of $\alpha{\downarrow}$ are axioms, where α is any variable. We therefore know $\Box\alpha{\downarrow}$ is an axiom. Hence, by GEN, $\forall\alpha\Box\alpha{\downarrow}$. Now let φ be $\Box\alpha{\downarrow}$. Then as an instance of axiom (39.1), we know:

$$\forall\alpha\Box\alpha{\downarrow} \to (\tau{\downarrow} \to \Box\tau{\downarrow}), \text{ where } \tau \text{ is } \textit{any} \text{ term substitutable for } \alpha \text{ in } \Box\alpha{\downarrow}$$

But every term τ of the same type as α is substitutable for α in $\Box\alpha{\downarrow}$. So it follows by Rule MP that $\tau{\downarrow} \to \Box\tau{\downarrow}$. $\bowtie$

(107.1) Assume $\tau = \sigma$. Then we prove $\tau{\downarrow}$ by cases:

(*A*) τ is an individual term.

(*B*) τ is an unary relation term.

(*C*) τ is an n-ary relation term ($n \geq 2$).

(*D*) τ is a 0-ary relation term.

Case A. τ is an individual term. Then τ and σ are both individual terms, say κ and κ', respectively, so that our assumption is $\kappa = \kappa'$. It then follows from this assumption, the definition of = (23.1), our conventions for definitions (17.2), and the Rule of Equivalence by Definition (90), that:

$$(O!\kappa \,\&\, O!\kappa' \,\&\, \forall F(F\kappa \equiv F\kappa')) \vee (A!\kappa \,\&\, A!\kappa' \,\&\, \Box\forall F(\kappa F \equiv \kappa' F)),$$
where F doesn't occur free in κ or κ'

So we reason by cases from the disjuncts. If $O!\kappa \,\&\, O!\kappa' \,\&\, \forall F(F\kappa \equiv F\kappa')$, then $O!\kappa$, and so by axiom (39.5.a), $\kappa\downarrow$, i.e., $\tau\downarrow$. If $A!\kappa \,\&\, A!\kappa' \,\&\, \Box\forall F(\kappa F \equiv \kappa' F)$, then $A!\kappa$ and, by axiom (39.5.a), $\kappa\downarrow$, i.e., $\tau\downarrow$.

Case B. τ is a unary relation term. Then τ and σ are both property terms, say Π and Π', so that our assumption is $\Pi = \Pi'$. Given (a) our conventions (17.2) for understanding the definition of property identity (23.2) and (b) Rule $\equiv_{df}$E of Definiendum Elimination (90.2), it follows that:

$$\Pi\downarrow \,\&\, \Pi'\downarrow \,\&\, \Box\forall x(x\Pi \equiv x\Pi'),$$
where x is any individual variable that doesn't occur free in Π or Π'.

Hence $\Pi\downarrow$, i.e., $\tau\downarrow$.

Case C. By reasoning analogous to Case B but with an appeal to the definition of relation identity (23.3).

Case D. Then $\tau\downarrow$ by (104.1). ⋈

(107.2) By reasoning analogous to (107.1). ⋈

(108.1) Let Π and Π' be any n-ary relation terms ($n \geq 0$) in which $x_1, \ldots, x_n$ don't occur free. Assume $\Pi = \Pi'$. Then by (107.1), it follows that $\Pi\downarrow$, and by (107.2), it follows that $\Pi'\downarrow$. Note independently that the following is axiomatic, since it is a closure of the axiom for the substitution of identicals (41):

$$\forall F\forall G(F = G \rightarrow$$
$$(\Box\forall x_1 \ldots \forall x_n(Fx_1 \ldots x_n \equiv Fx_1 \ldots x_n) \rightarrow \Box\forall x_1 \ldots \forall x_n(Fx_1 \ldots x_n \equiv Gx_x \ldots x_n)))$$

Since $x_1, \ldots, x_n$ don't occur free in Π and Π', the latter are substitutable, respectively, for F and G in the matrix of the above universal claim. So it follows by Rule $\forall$E that:

$$\Pi = \Pi' \rightarrow$$
$$(\Box\forall x_1 \ldots \forall x_n(\Pi x_1 \ldots x_n \equiv \Pi x_1 \ldots x_n) \rightarrow \Box\forall x_1 \ldots \forall x_n(\Pi x_1 \ldots x_n \equiv \Pi' x_x \ldots x_n))$$

And since $\Pi = \Pi'$ by assumption, it follows that:

$$\Box\forall x_1 \ldots \forall x_n(\Pi x_1 \ldots x_n \equiv \Pi x_1 \ldots x_n) \rightarrow \Box\forall x_1 \ldots \forall x_n(\Pi x_1 \ldots x_n \equiv \Pi' x_x \ldots x_n)$$

But the antecedent is clearly a theorem, by n applications of GEN and an application of RN to the instance $\Pi x_1 \ldots x_n \equiv \Pi x_1 \ldots x_n$ of the tautology $\varphi \equiv \varphi$. Hence $\Box\forall x_1 \ldots \forall x_n(\Pi x_1 \ldots x_n \equiv \Pi' x_x \ldots x_n)$. ⋈

(108.2) (Exercise)

(110) Assume $\Gamma_1 \vdash \varphi^\tau_\alpha$ and $\Gamma_2 \vdash \tau = \sigma$, where τ and σ are substitutable for α in φ. Now if we set Δ in (63.7) to $\Gamma_1 \cup \Gamma_2$, then we know:

(ξ) $\Gamma_1, \Gamma_2 \vdash \varphi^\tau_\alpha$

(ϑ) $\Gamma_1, \Gamma_2 \vdash \tau = \sigma$

We have to show $\Gamma_1, \Gamma_2 \vdash \varphi'$, where φ' is the result of replacing zero or more occurrences of τ in φ^τ_α with occurrences of σ. We begin by noting that theorems (107.1) and (107.2) respectively imply the following, by (63.10):

(a) $\tau = \sigma \vdash \tau\downarrow$

(b) $\tau = \sigma \vdash \sigma\downarrow$

Hence by applying (63.8) first to the pair (ϑ) and (a) and then to the pair (ϑ) and (b), we obtain, respectively:

(c) $\Gamma_1, \Gamma_2 \vdash \tau\downarrow$

(d) $\Gamma_1, \Gamma_2 \vdash \sigma\downarrow$

Since universal generalizations of the axiom schema for the substitution of identicals (41) are also axioms, we know by (63.1) that:

> $\vdash \forall\alpha\forall\beta(\alpha = \beta \to (\varphi \to \varphi''))$, whenever β is substitutable for α in φ and φ'' is the result of replacing zero or more free occurrences of α in φ with occurrences of β.

It follows from this last result by (63.3) that:

(e) $\Gamma_1, \Gamma_2 \vdash \forall\alpha\forall\beta(\alpha = \beta \to (\varphi \to \varphi''))$, whenever β is substitutable for α in φ and φ'' is the result of replacing zero or more free occurrences of α in φ with occurrences of β.

Now let ψ be $\forall\beta(\alpha = \beta \to (\varphi \to \varphi''))$ so that we may abbreviate (e) as:

(e) $\Gamma_1, \Gamma_2 \vdash \forall\alpha\psi$

Since τ is substitutable for α in φ, it is substitutable for α in ψ. So it follows from (e) and (c) by Rule $\forall$E (93.1) that $\Gamma_1, \Gamma_2 \vdash \psi^\tau_\alpha$, i.e.,

(f) $\Gamma_1, \Gamma_2 \vdash \forall\beta(\tau = \beta \to (\varphi^\tau_\alpha \to \varphi'''))$, where φ''' is the result of replacing zero or more occurrences of τ in φ^τ_α by occurrences of β.

Now abbreviate the proper scope of $\forall\beta$ in (f) as χ, so that we may abbreviate (f) as:

(f) $\Gamma_1, \Gamma_2 \vdash \forall\beta\chi$

Note that σ is substitutable for β in $\tau = \beta$, and is also substitutable for β in φ''' (exercise). So σ is substitutable for β in χ. Hence, it follows from (f) and (d) by Rule $\forall$E (93.1) that $\Gamma_1, \Gamma_2 \vdash \chi^{\sigma}_{\beta}$, i.e.,

$$(\tau = \beta \rightarrow (\varphi^{\tau}_{\alpha} \rightarrow \varphi'''))^{\sigma}_{\beta}$$

i.e.,

$$\tau = \sigma \rightarrow (\varphi^{\tau}_{\alpha} \rightarrow (\varphi''')^{\sigma}_{\beta})$$

Note that $(\varphi''')^{\sigma}_{\beta}$ is also the result of replacing zero or more occurrences of τ in φ^{τ}_{α} with occurrences of σ. That is, $(\varphi''')^{\sigma}_{\beta}$ is φ' and so we therefore know:

(g) $\Gamma_1, \Gamma_2 \vdash \tau = \sigma \rightarrow (\varphi^{\tau}_{\alpha} \rightarrow \varphi')$

Hence, from (g) and (ϑ), it follows by familiar reasoning that $\Gamma_1, \Gamma_2 \vdash \varphi^{\tau}_{\alpha} \rightarrow \varphi'$. From this last result and (ξ), it follows that $\Gamma_1, \Gamma_2 \vdash \varphi'$. ⋈

(111.1) The 0-ary instance of the axiom η-Conversion (48.3) asserts $[\lambda\, p] = p$, where p is a 0-ary relation variable. Since the closures of this principle are axioms, the following is therefore an axiom:

(ϑ) $\forall p([\lambda\, p] = p)$

Now by (104.2), we know $\varphi\downarrow$, for every formula φ. Moreover, since since λ binds no variables in $[\lambda\, p]$, every 0-ary relation term, and thus (4), every formula is substitutable for p in the matrix $[\lambda\, p] = p$ of (ϑ). So we may apply Rule $\forall$E (93.1) to instantiate φ for the quantifier $\forall p$ in (ϑ) to conclude $[\lambda\, \varphi] = \varphi$. ⋈

(111.2) $[\lambda\varphi] \equiv [\lambda\varphi]$ is an instance of the tautology $\psi \equiv \psi$. Moreover, by (111.1), we know $[\lambda\, \varphi] = \varphi$. So by Rule =E (110) it follows that $[\lambda\, \varphi] \equiv \varphi$. ⋈

(111.3) By axiom (39.2), $[\lambda\varphi]$ exists, i.e., $[\lambda\varphi]\downarrow$. From this and the 0-ary case of α-Conversion (48.1), it follows that $[\lambda\varphi] = [\lambda\varphi]'$, where $[\lambda\varphi]'$ is any alphabetic variant of $[\lambda\, \varphi]$. ⋈

(111.4) We know:

$$[\lambda\, \varphi] = [\lambda\, \varphi]' \qquad (111.3)$$

$$[\lambda\, \varphi] = \varphi \qquad (111.1)$$

So by Rule =E (110), it follows that $\varphi = [\lambda\, \varphi]'$. Independently, by the definition of alphabetic variants (16), we know $[\lambda\, \varphi]' = [\lambda\, \varphi']$. So by metalinguistic substitution of identicals, $\varphi = [\lambda\, \varphi']$. But as an instance of (111.1), we know $[\lambda\, \varphi'] = \varphi'$. Hence, again, by Rule =E, $\varphi = \varphi'$. ⋈

(111.5) (Exercise)

(111.6) ($\rightarrow$) Assume $\varphi \equiv \psi$. By (111.2), we know $[\lambda\, \varphi] \equiv \varphi$. Hence, by biconditional syllogism, $[\lambda\, \varphi] \equiv \psi$. But also by (111.2) and the commutativity of the biconditional, we know $\psi \equiv [\lambda\, \psi]$. Hence again by biconditional syllogism, $[\lambda\, \varphi] \equiv [\lambda\, \psi]$. ($\leftarrow$) Reverse the reasoning, starting with the assumption that $[\lambda\, \varphi] \equiv [\lambda\, \psi]$. ⋈

(114) By (111.5), we know $\vdash (\varphi \equiv \varphi')$, where φ' is any alphabetic variant of φ. Hence, by (63.3), $\Gamma \vdash (\varphi \equiv \varphi')$. By definition of $\equiv$ (18.3) and Rule $\equiv_{df}$E of Definiendum Elimination (90.2), it follows that $\Gamma \vdash ((\varphi \rightarrow \varphi') \,\&\, (\varphi' \rightarrow \varphi))$. So by Rule &E, it follows that:

(A) $\Gamma \vdash (\varphi \rightarrow \varphi')$

(B) $\Gamma \vdash (\varphi' \rightarrow \varphi)$

Now to justify the left-to-right direction of the Rule of Alphabetic Variants, assume $\Gamma \vdash \varphi$. Then from this and (A), it follows by (63.5) that $\Gamma \vdash \varphi'$. By analogous reasoning from (B), if $\Gamma \vdash \varphi'$, then $\Gamma \vdash \varphi$. ⋈

(115.1) By definition (22.1), i.e., $O! =_{df} [\lambda x\, \Diamond E!x]$, the Rule of Definition by Identity (73) tells us that:

$$([\lambda x\, \Diamond E!x]\downarrow \rightarrow (O! = [\lambda x\, \Diamond E!x])) \,\&\, (\neg[\lambda x\, \Diamond E!x]\downarrow \rightarrow \neg O!\downarrow)$$

So by &E, it follows that:

$$[\lambda x\, \Diamond E!x]\downarrow \rightarrow (O! = [\lambda x\, \Diamond E!x])$$

But since $[\lambda x\, \Diamond E!x]$ is a core λ-expression, as this notion was defined in (9.2), axiom (39.2) asserts $[\lambda x\, \Diamond E!x]\downarrow$. Hence, $O! = [\lambda x\, \Diamond E!x]$. So by (107.1), $O!\downarrow$. ⋈

(115.2) (Exercise)

(115.3) Since we haven't yet established the symmetry of identity, we reason as follows.[428] By β-Conversion and the fact that $[\lambda x\, \Diamond E!x]\downarrow$ (39.2), we know:

(ϑ) $[\lambda x\, \Diamond E!x]x \equiv \Diamond E!x$

[428] Without the symmetry of identity, we can't yet use the fact that $O! = [\lambda x\, \Diamond E!x]$, which we established as part of the proof of (115.1), to infer $[\lambda x \Diamond E!x] = O!$. If we could establish the latter, we could then infer $O!x \equiv \Diamond E!x$ by Rule =E by from an instance of β-Conversion (namely, $[\lambda x \Diamond E!x]\downarrow \rightarrow ([\lambda x\, \Diamond E!x]x \equiv \Diamond E!x)$) and the fact that $[\lambda x\, \Diamond E!x]\downarrow$ is axiomatic (39.2). So, a less direct proof has to be found.

Independently, as an instance of (88.3.a), we know:

(ζ) $O!x \equiv O!x$

But in the proof of (115.1), we established that $O! = [\lambda x \Diamond E!x]$. By applying this last fact and Rule =E to (ζ), we obtain:

(ξ) $O!x \equiv [\lambda x \Diamond E!x]x$

Then by biconditional syllogism (89.3.e), (ξ) and (ϑ) imply $O!x \equiv \Diamond E!x$. ⋈

(115.4) (Exercise)

(115.5) By (83), $\Diamond E!x \vee \neg\Diamond E!x$. But by the commutativity of the biconditional (88.2.e), theorems (115.3) and (115.4) imply, respectively:

$\Diamond E!x \equiv O!x$

$\neg\Diamond E!x \equiv A!x$

So by disjunctive syllogism (89.1), it follows that $O!x \vee A!x$. ⋈

(116.1) By definition of property identity (23.2) and Rule of Equivalence by Definition (90.1), we know:

$F = G \equiv (F{\downarrow} \;\&\; G{\downarrow} \;\&\; \Box\forall x(xF \equiv xG))$

But by axiom (39.2), we know both $F{\downarrow}$ and $G{\downarrow}$. Hence by the general form of the Rule ≡S of Biconditional Simplification (91), it follows that:

$F = G \equiv \Box\forall x(xF \equiv xG)$ ⋈

(116.2) – **(116.3)** (Exercises)

(117.1) We prove this by cases, using the four cases of definition of identity (23) as our guide.[429] The four cases we have to show are:

[429] In a personal communication (19 January 2021), Daniel Kirchner notes that his implementation of object theory in Isabelle/HOL produced the following simpler proof, where the last 3 cases are collapsed into a single case. His system showed that $F = F$, where F is any n-ary relation variable ($n \geq 0$), as follows. Fix n. By the axiom for η-Conversion (48.3), we know:

(ϑ) $[\lambda x_1 \ldots x_n \, Fx_1 \ldots x_n] = F$

Now the following is an instance of axiom (39.2):

$[\lambda x_1 \ldots x_n \, Fx_1 \ldots x_n]{\downarrow}$

and since $[\lambda x_1 \ldots x_n \, Fx_1 \ldots x_n]$ is an alphabetic variant of itself, the following is an instance of the axiom for α-Conversion (48.1):

$[\lambda x_1 \ldots x_n \, Fx_1 \ldots x_n]{\downarrow} \to [\lambda x_1 \ldots x_n \, Fx_1 \ldots x_n] = [\lambda x_1 \ldots x_n \, Fx_1 \ldots x_n]$

Our last two results therefore imply:

(ξ) $[\lambda x_1 \ldots x_n \, Fx_1 \ldots x_n] = [\lambda x_1 \ldots x_n \, Fx_1 \ldots x_n]$

- $x = x$
- $F = F$, where F is a unary relation variable
- $F = F$, where F is an n-ary relation variable, $n \geq 2$
- $p = p$, where p is a 0-ary relation variable

Case 1. By definition (23.1), we have to show:

(ϑ) $(O!x \,\&\, O!x \,\&\, \Box\forall F(Fx \equiv Fx)) \vee (A!x \,\&\, A!x \,\&\, \Box\forall F(xF \equiv xF))$

To establish (ϑ), our proof strategy is to reason by disjunctive syllogism (86.3.c) as follows:

From:

(a) $O!x \vee A!x$

(b) $O!x \rightarrow (O!x \,\&\, O!x \,\&\, \Box\forall F(Fx \equiv Fx))$

(c) $A!x \rightarrow (A!x \,\&\, A!x \,\&\, \Box\forall F(xF \equiv xF))$

conclude (ϑ).

So if we can establish (a), (b), and (c), we're done. But (a) is just theorem (115.5). So it remains to show (b) and (c):[430]

(b) Assume $O!x$. By the idempotency of & (85.6), it follows that $O!x \,\&\, O!x$. Note, independently that as an instance of (88.3.a), we have $Fx \equiv Fx$. Since this is a theorem, we may apply GEN (66) to obtain $\forall F(Fx \equiv Fx)$. Since this is a $\Box$-theorem, it follows by RN (68) that $\Box\forall F(Fx \equiv Fx)$. If we conjoin this last result by &I (86.1) with what we have established so far, we have $O!x \,\&\, O!x \,\&\, \Box\forall F(Fx \equiv Fx)$.

(c) Assume $A!x$. By the idempotency of & (85.6), it follows that $A!x \,\&\, A!x$. Note, independently that as an instance of (88.3.a), we have $xF \equiv xF$. Since this is a theorem, we may apply GEN to obtain $\forall F(xF \equiv xF)$. Since this is a $\Box$-theorem, it follows by RN that $\Box\forall F(xF \equiv xF)$. If we conjoin this last result by &I (86.1) with what we have established so far, we have $A!x \,\&\, A!x \,\&\, \Box\forall F(xF \equiv xF)$.

But the Variant of Rule =E (110) tells us that $\varphi^{\tau}_{\alpha}, \tau = \sigma \vdash \varphi'$, where φ' is the result of replacing zero or more occurrences of τ in φ^{τ}_{α} with occurrences of σ. So if we let φ be $\alpha = \alpha$, let τ be $[\lambda x_1 \ldots x_n\, Fx_1 \ldots x_n]$, and let σ be F, the following is an instance:

$$[\lambda x_1 \ldots x_n\, Fx_1 \ldots x_n] = [\lambda x_1 \ldots x_n\, Fx_1 \ldots x_n],\ [\lambda x_1 \ldots x_n\, Fx_1 \ldots x_n] = F \vdash F = F \qquad \text{(Rule =E)}$$

So from (ϑ) and (ξ) it follows by this instance of Rule =E that $F = F$. ⋈

[430] Thanks go to Wes Anderson, who correctly reported that in an earlier draft, the proof universally generalized, incorrectly, on the variable x instead of the variable F, in (b) and (c).

Case 2. Note that if we apply GEN *once* to theorem (116.1), we obtain:

$$\forall G(F = G \equiv \Box\forall x(xF \equiv xG))$$

But since $F\downarrow$ and F is substitutable for G in the matrix of the above, we may instantiate the above result to F, by (93.2), to obtain:

$$F = F \equiv \Box\forall x(xF \equiv xF)$$

Hence, to prove $F = F$, it suffices to show $\Box\forall x(xF \equiv xF)$. But, as an instance of (88.3.a), we know $xF \equiv xF$. Since this is a theorem, we may apply GEN (66) to obtain $\forall x(xF \equiv xF)$. Since this is a $\Box$-theorem, it follows by RN (68) that $\Box\forall x(xF \equiv xF)$.

Case 3. If we apply GEN once to theorem (116.2), then we know:

$$\begin{aligned}
&\forall G(F^n = G^n \equiv \\
&\quad \forall y_1 \ldots \forall y_{n-1}([\lambda x\, F^n xy_1 \ldots y_{n-1}] = [\lambda x\, G^n xy_1 \ldots y_{n-1}]\ \&\\
&\qquad [\lambda x\, F^n y_1 xy_2 \ldots y_{n-1}] = [\lambda x\, G^n y_1 xy_2 \ldots y_{n-1}]\ \& \ldots \&\\
&\qquad [\lambda x\, F^n y_1 \ldots y_{n-1} x] = [\lambda x\, G^n y_1 \ldots y_{n-1} x]))
\end{aligned}$$

Since $F\downarrow$ and F is a variable and substitutable for G in the matrix of the above, we may instantiate the above result to F, by (93.2), to obtain:

$$\begin{aligned}
&F^n = F^n \equiv \\
&\quad \forall y_1 \ldots \forall y_{n-1}([\lambda x\, F^n xy_1 \ldots y_{n-1}] = [\lambda x\, F^n xy_1 \ldots y_{n-1}]\ \&\\
&\qquad [\lambda x\, F^n y_1 xy_2 \ldots y_{n-1}] = [\lambda x\, F^n y_1 xy_2 \ldots y_{n-1}]\ \& \ldots \&\\
&\qquad [\lambda x\, F^n y_1 \ldots y_{n-1} x] = [\lambda x\, F^n y_1 \ldots y_{n-1} x])
\end{aligned}$$

So to prove $F^n = F^n$, it suffices to show:

$$\begin{aligned}
&\forall y_1 \ldots \forall y_{n-1}([\lambda x\, F^n xy_1 \ldots y_{n-1}] = [\lambda x\, F^n xy_1 \ldots y_{n-1}]\ \&\\
&\qquad [\lambda x\, F^n y_1 xy_2 \ldots y_{n-1}] = [\lambda x\, F^n y_1 xy_2 \ldots y_{n-1}]\ \& \ldots \&\\
&\qquad [\lambda x\, F^n y_1 \ldots y_{n-1} x] = [\lambda x\, F^n y_1 \ldots y_{n-1} x])
\end{aligned}$$

By $n-1$ applications of GEN, it suffices to show:

$$\begin{aligned}
&[\lambda x\, F^n xy_1 \ldots y_{n-1}] = [\lambda x\, F^n xy_1 \ldots y_{n-1}]\ \&\\
&\qquad [\lambda x\, F^n y_1 xy_2 \ldots y_{n-1}] = [\lambda x\, F^n y_1 xy_2 \ldots y_{n-1}]\ \& \ldots \&\\
&\qquad [\lambda x\, F^n y_1 \ldots y_{n-1} x] = [\lambda x\, F^n y_1 \ldots y_{n-1} x]
\end{aligned}$$

But note that all of the λ-expressions in the above are core λ-expressions. Hence, by (39.2), we know:

$$\begin{aligned}
&[\lambda x\, F^n xy_1 \ldots y_{n-1}]\downarrow\\
&[\lambda x\, F^n y_1 xy_2 \ldots y_{n-1}]\downarrow\\
&\quad\vdots\\
&[\lambda x\, F^n y_1 \ldots y_{n-1} x]\downarrow
\end{aligned}$$

So we may instantiate, one-by-one, each of the above terms into the universal generalization of *Case 2* and we obtain each of the conjuncts of what we had to show.

Case 4. By applying GEN to *Case 2* above, it follows that $\forall F(F=F)$, where F is a property variable. Note that $[\lambda x\, p]\downarrow$ (39.2) and $[\lambda x\, p]$ is substitutable for F in the matrix $F=F$. So we may use Rule $\forall$E (93.2) to infer $[\lambda x\ p]=[\lambda x\ p]$. Now by a single application of GEN to theorem (116.3), we know $\forall q(p=q \equiv [\lambda x\, p]=[\lambda x\, q])$. Since p is substitutable for q in the matrix of this universal claim and p is a variable, it follows, by Rule $\forall$E (93.2), that $p=p \equiv [\lambda x\, p]=[\lambda x\, p]$. Since we've established $[\lambda x\, p]=[\lambda x\, p]$, it follows by biconditional syllogism (89.3.b) that $p = p$. ⋈

(117.2) By (117.1), we know $\alpha=\alpha$. For conditional proof, assume $\alpha=\beta$. Hence by the Variant version of Rule =E (110), it follows that $\beta=\alpha$.[431] ⋈

(117.3) Assume the antecedent, so that by &E we know both $\alpha=\beta$ and $\beta=\gamma$. Then by Rule =E (110) it follows that $\alpha=\gamma$. ⋈

(117.4) ($\rightarrow$) Note that for any distinct variable γ of the same type as α, it is an easy exercise to show $\forall\gamma(\alpha=\gamma \equiv \alpha=\gamma)$. So for conditional proof, assume $\alpha=\beta$. Then by Rule =E (110), $\forall\gamma(\alpha=\gamma \equiv \beta=\gamma)$. ($\leftarrow$) Assume $\forall\gamma(\alpha=\gamma \equiv \beta=\gamma)$. Then since $\alpha\downarrow$ and is substitutable for γ in the matrix of this universal claim, it follows by by Rule $\forall$E (93.2) that $\alpha=\alpha \equiv \beta=\alpha$. But $\alpha=\alpha$, by (117.1). So $\beta=\alpha$, by biconditional syllogism (89.3.a), and thus $\alpha=\beta$, by (117.2). ⋈

(118.1) Assume:

(ϑ) $\Gamma \vdash \tau\downarrow$

Now let us independently establish that:

(ξ) $\vdash \tau\downarrow \rightarrow \tau=\tau$

> *Proof*. To see that there is a proof of $\tau\downarrow \rightarrow \tau=\tau$, assume for conditional proof that $\tau\downarrow$. Independently, by (117.1), it is a theorem that $\alpha=\alpha$. So this theorem yields $\forall\alpha(\alpha = \alpha)$ by GEN. Now since α occurs free in $\alpha=\alpha$ and doesn't occur within the scope of any variable-binding operators, τ is substitutable for α in $\alpha=\alpha$. Hence, by Rule $\forall$E (93.1) it follows that $\tau=\tau$.

Then from (ϑ) and (ξ) it follows that $\Gamma \vdash \tau=\tau$, by (63.5). ⋈

[431] To obtain the relevant instance of the rule we just used, set the metavariables in Rule =E to the following values: let τ be α, σ be β, α be γ, φ be $\gamma=\gamma$ (so that φ^{τ}_{α} is $\alpha=\alpha$), and φ' be $\beta=\alpha$ (thus, φ' is the result of replacing one occurrence of α in φ^{τ}_{α} by β). Given these assignments, the rule asserts: $\alpha=\alpha,\, \alpha=\beta \vdash \beta=\alpha$.

(118.2) Assume τ is a primitive constant, a variable, or a core λ-expression. Then by (39.2), $\tau\downarrow$ is an axiom and, hence, a theorem. Since this establishes $\vdash \tau\downarrow$, it follows by Rule =I (118.1) that $\vdash \tau = \tau$. ⋈

(120.1) By hypothesis:

(A) $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$ is a definition-by-= in which the variables $\alpha_1,\ldots,\alpha_n$ occur free ($n \geq 0$),

(B) $\tau_1,\ldots,\tau_n$ are substitutable, respectively, for $\alpha_1,\ldots,\alpha_n$ in both definiens and definiendum, and

(C) $\Gamma \vdash \sigma(\tau_1,\ldots,\tau_n)\downarrow$.

From (A), (B), and the Rule of Definition by Identity (73), we know:

$\vdash (\sigma(\tau_1,\ldots,\tau_n)\downarrow \rightarrow \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n))\ \&\ (\neg\sigma(\tau_1,\ldots,\tau_n)\downarrow \rightarrow \neg\tau(\tau_1,\ldots,\tau_n)\downarrow)$

By (63.3), the above holds for any premise set Γ:

(D) $\Gamma \vdash (\sigma(\tau_1,\ldots,\tau_n)\downarrow \rightarrow \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n))\ \&\ (\neg\sigma(\tau_1,\ldots,\tau_n)\downarrow \rightarrow \neg\tau(\tau_1,\ldots,\tau_n)\downarrow)$

By Rule &E (86.2.a), it follows from (D) that:

(E) $\Gamma \vdash \sigma(\tau_1,\ldots,\tau_n)\downarrow) \ \rightarrow \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n)$

Hence, from (E) and (C) it follows by (63.5) that:

$$\Gamma \vdash \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n) \qquad ⋈$$

(120.2.a) By hypothesis:

(A) $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$ is a definition-by-= in which $\alpha_1,\ldots,\alpha_n$ occur free ($n \geq 0$),

(B) $\tau_1,\ldots,\tau_n$ are substitutable for $\alpha_1,\ldots,\alpha_n$, respectively, in both definiens and definiendum,

(C) φ contains one or more occurrences of $\tau(\tau_1,\ldots,\tau_n)$, and

(D) φ' is the result of replacing zero or more occurrences of $\tau(\tau_1,\ldots,\tau_n)$ in φ by $\sigma(\tau_1,\ldots,\tau_n)$

Now assume:

(E) $\Gamma \vdash \sigma(\tau_1,\ldots,\tau_n)\downarrow$

(F) $\Gamma \vdash \varphi$

(A), (B), and (E) imply, by the Rule of Identity by Definition (120.1), that:

(G) $\Gamma \vdash \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n)$

Then by (C), (D), and Rule =E (110), it follows from (G) and (F) that $\Gamma \vdash \varphi'$. ⋈

(120.2.b) By hypothesis:

(A) $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$ is a definition-by-= in which $\alpha_1,\ldots,\alpha_n$ occur free ($n \geq 0$),

(B) $\tau_1,\ldots,\tau_n$ are substitutable for $\alpha_1,\ldots,\alpha_n$, respectively, in both definiens and definiendum,

(C) φ contains one or more occurrences of $\sigma(\tau_1,\ldots,\tau_n)$, and

(D) φ' is the result of replacing zero or more occurrences of $\sigma(\tau_1,\ldots,\tau_n)$ in φ by $\tau(\tau_1,\ldots,\tau_n)$

Now assume:

(E) $\Gamma \vdash \sigma(\tau_1,\ldots,\tau_n)\downarrow$

(F) $\Gamma \vdash \varphi$

(A), (B), and (E) imply, by the Rule of Identity by Definition (120.1), that:

(G) $\Gamma \vdash \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n)$

Now note independently that since (107.1) is a theorem, we know:

$$\vdash \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n) \to \tau(\tau_1,\ldots,\tau_n)\downarrow$$

It follows from this last fact by (63.10) that:

(H) $\tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n) \vdash \tau(\tau_1,\ldots,\tau_n)\downarrow$

Hence, from (G), and (H), it follows by (63.8) that the significance of the definiendum is derivable from Γ, i.e., that:

(I) $\Gamma \vdash \tau(\tau_1,\ldots,\tau_n)\downarrow$

But note independently that it is a theorem that identity is symmetric (117.2), i.e., that $\vdash \alpha = \beta \to \beta = \alpha$. So two applications of GEN:

(J) $\vdash \forall\alpha\forall\beta(\alpha = \beta \to \beta = \alpha)$

From (J), it follows by (63.3) that:

(K) $\Gamma \vdash \forall\alpha\forall\beta(\alpha = \beta \to \beta = \alpha)$

But by (E) and (I), we know that the significance of the terms $\tau(\tau_1,\ldots,\tau_n)$ and $\sigma(\tau_1,\ldots,\tau_n)$ is derivable from Γ. So by two applications of Rule $\forall$E (93.1), it follows from (K) that:

(L) $\Gamma \vdash \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n) \rightarrow \sigma(\tau_1,\ldots,\tau_n) = \tau(\tau_1,\ldots,\tau_n)$

So from (L) and (G), it follows by (63.5) that:

(M) $\Gamma \vdash \sigma(\tau_1,\ldots,\tau_n) = \tau(\tau_1,\ldots,\tau_n)$

Then by Rule =E (110), it follows from (F) and (M) that:

$\Gamma \vdash \varphi'$ ⋈

(121.1) Consider any term τ in which β doesn't occur free. ($\rightarrow$) Assume $\tau\downarrow$. Then by Rule =I (118.1), $\tau = \tau$. Now let φ be $\beta = \tau$. Note that since β doesn't occur free in τ, φ has just one free occurrence of β, namely, the initial occurrence. So $\tau = \tau$ has the form φ^τ_β. Note also that τ is substitutable for β in $\beta = \tau$, that is, τ is substitutable for β in φ. And since $\tau\downarrow$, all the conditions for applying Rule $\exists$I (101.1) are met. So we can infer $\exists\beta\varphi$ from φ^τ_β, i.e., infer $\exists\beta(\beta = \tau)$ from $\tau = \tau$.
($\leftarrow$) Assume $\exists\beta(\beta = \tau)$. Now suppose σ is an arbitrary such entity, so that we know $\sigma = \tau$; formally, σ is a *fresh*, primitive constant of the same type as β that hasn't previously appeared in $\beta = \tau$, or $\tau\downarrow$, or in any axioms in $\boldsymbol{\Lambda}$.[432] Then by (107.2), it follows that $\tau\downarrow$. Since we have reached $\tau\downarrow$ from $\sigma = \tau$ and the conditions of Rule $\exists$E (102) are met, we may use Rule $\exists$E to conclude that we have derived $\tau\downarrow$ from our assumption $\exists\beta(\beta = \tau)$. ⋈

(121.2) By hypothesis, τ is substitutable for α in φ and β doesn't occur free in τ. Now, for conditional proof, assume $\forall\alpha\varphi$ to show $\exists\beta(\beta = \tau) \rightarrow \varphi^\tau_\alpha$. And for a nested conditional proof, assume $\exists\beta(\beta = \tau)$, to show φ^τ_α. From our second assumption (given our hypothesis that β doesn't occur free in τ), it follows from the right-to-left direction of the previous theorem (121.1) that $\tau\downarrow$. Hence, from our first assumption and this last result (given our hypothesis that τ is substitutable for α in φ), it follows by axiom (39.1) that φ^τ_α. ⋈

(121.3) From (121.1) and axiom (39.2). ⋈

(121.4) From (121.1) and axioms (39.5.a) and (39.5.b). ⋈

(123.1.a) Since α is a variable, $\alpha\downarrow$ is an instance of axiom (39.2) and, hence, a theorem. By GEN, $\forall\alpha\,\alpha\downarrow$. ⋈

(123.1.b) [There are several different ways to prove this theorem. Here is one.] Let α and β be distinct variables of the same type. Since α is a variable, $\alpha\downarrow$ is an instance of axiom (39.2). Since β doesn't occur free in α, it follows by the

[432] Without the proviso that β doesn't occur free in τ, the assumption that σ is an arbitrary witness to $\exists\beta(\beta = \tau)$ yields only the knowledge that $\sigma = \tau^\sigma_\beta$. And this fact doesn't allow us to conclude $\tau\downarrow$; it only yields $\tau^\sigma_\beta\downarrow$. But with the proviso that β doesn't occur free in τ, the term τ^σ_β is just τ itself. So $\sigma = \tau^\sigma_\beta$ is the same expression as $\sigma = \tau$. This explains why the proviso that β doesn't occur free in τ is essential to the right-to-left direction of this theorem.

left-to-right direction of (121.1) that $\exists\beta(\beta = \alpha)$. Hence, by GEN, $\forall\alpha\exists\beta(\beta = \alpha)$. ⋈

(123.2.a) Since α is a variable, $\alpha\downarrow$ is an instance of (39.2). So, by RN (68), $\Box\alpha\downarrow$. ⋈

(123.2.b) From (123.1.b), it follows that $\exists\beta(\beta = \alpha)$, by Rule ∀E (93.3). Since the derivation was modally strict, it follows by RN (68) that $\Box\exists\beta(\beta = \alpha)$. ⋈

(123.3.a) Since (123.1.a) is a □-theorem, we may apply RN (68) to obtain $\Box\forall\alpha\,\alpha\downarrow$. ⋈

(123.3.b) Since (123.1.b) is a □-theorem, we may apply RN (68) to obtain $\Box\forall\alpha\exists\beta(\beta = \alpha)$. ⋈

(123.4.a) From (123.2.a), by GEN. ⋈

(123.4.b) From (123.2.b), by GEN. ⋈

(123.5.a) From the □-theorem (123.4.a), by RN (68). ⋈

(123.5.b) From the □-theorem (123.4.b), by RN (68). ⋈

(124.1) Apply GEN to theorem (117.1). Since the result is a □-theorem, apply RN. ⋈

(124.2) Since (117.1) is a □-theorem, we may apply RN to obtain $\Box(\alpha = \alpha)$ as a theorem. Hence, by GEN, we have: $\forall\alpha\Box(\alpha = \alpha)$. ⋈

(125.1) Assume $\alpha = \beta$, for conditional proof. Since $\alpha = \alpha$ (117.1) is a □-theorem, we know by RN (68) that $\Box\alpha = \alpha$. Then by Rule =E (110), it follows that $\Box\alpha = \beta$. ⋈

(125.2) Assume $\tau = \sigma$, for conditional proof. Then by (107.1) and (107.2), we know both $\tau\downarrow$ and $\sigma\downarrow$. Now since (125.1) is a theorem, it follows by two applications of GEN that:

$$\forall\alpha\forall\beta(\alpha = \beta \to \Box\alpha = \beta)$$

Hence, by two applications of Rule ∀E (93.1), it follows that $\tau = \sigma \to \Box\tau = \sigma$. ⋈

(126.1) Suppose β is substitutable for α in φ and doesn't occur free in φ. Then there are two cases.

Case 1. β just is α. Then our theorem states $\varphi \equiv \exists\alpha(\alpha{=}\alpha \;\&\; \varphi^{\alpha}_{\alpha})$. By hypothesis, β, i.e., α, doesn't occur free in φ and so $\varphi^{\alpha}_{\alpha} = \varphi$. So our theorem states $\varphi \equiv \exists\alpha(\alpha{=}\alpha \;\&\; \varphi)$. We leave the remainder of the proof as an exercise.

Case 2. β and α are distinct variables of the same type. (→) Assume φ, for conditional proof. By definition of substitutions, the formula φ is identical to the formula $\varphi^{\alpha}_{\alpha}$. Since it is a theorem (117.1) that $\alpha = \alpha$, we have, by &I, that $\alpha = \alpha \;\&\; \varphi^{\alpha}_{\alpha}$. Hence, by ∃I, it follows that $\exists\beta(\beta = \alpha \;\&\; \varphi^{\beta}_{\alpha})$. (←) Assume, for conditional proof:

(ϑ) $\exists\beta(\beta=\alpha \,\&\, \varphi^{\beta}_{\alpha})$,

Assume τ is an arbitrary such entity; formally, the metavariable τ stands for some new, primitive constant that doesn't appear in (ϑ) or φ, that has the same type as the variable that β takes as value, and that is therefore substitutable for β in the matrix of (ζ). So we know:

(ζ) $\tau=\alpha \,\&\, (\varphi^{\beta}_{\alpha})^{\tau}_{\beta}$

We may detach the two conjuncts of (ζ) by &E. Since β is substitutable for α in φ and doesn't occur free in φ, and τ is substitutable for α in φ, the Re-replacement theorem (97.3) tells us that $(\varphi^{\beta}_{\alpha})^{\tau}_{\beta} = \varphi^{\tau}_{\alpha}$. So it follows from the right conjunct of (ζ) that φ^{τ}_{α}. But from this latter conclusion and the left conjunct of (ζ), it follows by Rule =E (110) that we may substitute α for every occurrence of τ in φ^{τ}_{α}, to obtain $(\varphi^{\tau}_{\alpha})^{\alpha}_{\tau}$. But, by Re-replacement lemma (97.2), since τ is a primitive constant that doesn't appear in φ, this is just $\varphi^{\alpha}_{\alpha}$, i.e., φ. By $\exists$E (102), we can discharge (ζ) and conclude φ. ⋈

(126.2) Assume $\tau\downarrow$, where τ is, by hypothesis, substitutable for α in φ. Now choose any variable of the same type as α, say β, that is substitutable for α in φ and that doesn't occur free in φ. Then by applying GEN to (126.1), we know the following applies to φ:

(ϑ) $\forall\alpha(\varphi \equiv \exists\beta(\beta=\alpha \;\&\; \varphi^{\beta}_{\alpha}))$

Since τ is substitutable for α in φ, it is substitutable for α in the matrix of (ϑ). So since $\tau\downarrow$, we may instantiate τ into $\forall\alpha$ in (ϑ) by $\forall$E, to obtain:

$$\varphi^{\tau}_{\alpha} \equiv \exists\beta(\beta=\tau \;\&\; \varphi^{\beta}_{\alpha}))$$

But given our choice of β, we know that by commuting an appropriate instance of (103.7), the following applies to φ:

$$\exists\beta(\beta=\tau \,\&\, \varphi^{\beta}_{\alpha}) \equiv \exists\alpha(\alpha=\tau \,\&\, \varphi)$$

By biconditional syllogism (89.3.e), it follows that:

$$\varphi^{\tau}_{\alpha} \equiv \exists\alpha(\alpha=\tau \,\&\, \varphi) \qquad ⋈$$

(126.3) Suppose α,β are distinct variables and consider any formula φ in which β is substitutable for α and doesn't occur free. ($\rightarrow$) Assume:

(ζ) $\varphi \,\&\, \forall\beta(\varphi^{\beta}_{\alpha} \rightarrow \beta=\alpha)$

By theorem (99.2), it suffices to show $\forall\beta(\varphi^{\beta}_{\alpha} \rightarrow \beta=\alpha) \,\&\, \forall\beta(\beta=\alpha \rightarrow \varphi^{\beta}_{\alpha})$. Since the 1st conjunct is the 2nd conjunct of (ζ), it remains only to show $\forall\beta(\beta=\alpha \rightarrow \varphi^{\beta}_{\alpha})$. By hypothesis, β doesn't occur free in φ and, hence, doesn't occur free in our assumption (ζ). So by GEN, it remains to show $\beta = \alpha \to \varphi^{\beta}_{\alpha}$. So assume $\beta = \alpha$, which by the symmetry of identity (117.2), yields $\alpha = \beta$. Now (ζ) also implies φ. Since β is, by hypothesis, substitutable for α in φ, it follows by Rule =E (110) that φ^{β}_{α}. ($\leftarrow$) Assume:

(ϑ) $\forall\beta(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha)$

for conditional proof. By &I, it suffices to show:

(a) φ

(b) $\forall\beta(\varphi^{\beta}_{\alpha} \to \beta = \alpha)$

(a) Since, by hypothesis, β is substitutable for α in φ and isn't free in φ, it follows by the Re-replacement lemma (97.1) that α is substitutable for β in φ^{β}_{α}. Hence α is substitutable for β in $\varphi^{\beta}_{\alpha} \equiv \beta = \alpha$. So we may, by Rule $\forall$E (93.2), instantiate $\forall\beta$ in (ϑ) to α, and thereby obtain:

$$(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha)^{\alpha}_{\beta}$$

By the definition of ψ^{α}_{β} (14) extended to include defined formulas of the form $\psi \equiv \chi$, this becomes:

(ξ) $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta} \equiv (\beta = \alpha)^{\alpha}_{\beta}$

Since the Re-replacement lemma is operative, the left condition of (ξ) is just φ. By definition of substitutions, the right condition of (ξ) is $\alpha = \alpha$ (this is obvious, but we leave the strict proof, by way of the cases in the definition of =, as an exercise). Hence, (ξ) resolves to:

$$\varphi \equiv \alpha = \alpha$$

But since we know $\alpha = \alpha$ by (117.1), it follows that φ, by biconditional syllogism. (b) By (99.2), it follows from (ϑ) that:

$$\forall\beta(\varphi^{\beta}_{\alpha} \to \beta = \alpha) \;\&\; \forall\beta(\beta = \alpha \to \varphi^{\beta}_{\alpha})$$

So $\forall\beta(\varphi^{\beta}_{\alpha} \to \beta = \alpha)$ follows by &E. ⋈

(126.4) Suppose α, β are distinct variables of the same type and consider any formula φ in which β is substitutable for α in φ and doesn't occur free in φ. We show both directions of the biconditional.

($\to$) Assume:

(ζ) $\varphi^{\beta}_{\alpha} \;\&\; \forall\alpha(\varphi \to \alpha = \beta)$

Since α doesn't occur free in (ζ), it suffices by GEN to show $\varphi \equiv \alpha = \beta$. ($\rightarrow$) Assume φ. Note that from the second conjunct of (ζ), it follows by the special case of Rule $\forall$E (93.3) that $\varphi \rightarrow \alpha = \beta$. Hence $\alpha = \beta$. ($\leftarrow$) Assume $\alpha = \beta$. From this and the first conjunct of (ζ) it follows by the Rule =E (110) that $(\varphi^{\beta}_{\alpha})'$, where $(\varphi^{\beta}_{\alpha})'$ is the result of replacing zero or more occurrences of β in φ^{β}_{α} by α. Since this includes the special case where we replace *all* the occurrences of β in φ^{β}_{α} by α, i.e., $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$. But by hypothesis, β is substitutable for α in φ and doesn't occur free in φ. So the Re-replacement Lemma (97.1) applies and $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$ just is φ.

($\leftarrow$) Assume $\forall\alpha(\varphi \equiv \alpha = \beta)$. Then since β is substitutable for α in φ, we can instantiate β into the universal claim to obtain $\varphi^{\beta}_{\alpha} \equiv \beta = \beta$. Since $\beta = \beta$ is a theorem, it follows that φ^{β}_{α}. So it remains to show $\forall\alpha(\varphi \rightarrow \alpha = \beta)$, which is left as an exercise. ⋈

(127.2) (Exercise)

(128) Assume $\exists!\alpha\varphi$, where (by hypothesis) β and γ are variables that don't occur free, and are substitutable for α, in φ. Then by a fact about uniqueness quantifier (127.2) and Rule $\equiv_{df}$E of Definiendum Elimination (90.2), $\exists\alpha\forall\beta(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha)$. Suppose τ is such an α, i.e., let τ be an arbitrarily chosen, primitive constant of the same type as α and assume:

(ϑ) $\forall\beta(\varphi^{\beta}_{\alpha} \equiv \beta = \tau)$

By hypothesis, γ also doesn't occur free, and is substitutable for α, in φ. So we want to show:

$$\forall\beta\forall\gamma((\varphi^{\beta}_{\alpha} \,\&\, \varphi^{\gamma}_{\alpha}) \rightarrow \beta = \gamma)$$

Since neither β nor γ occur free in any assumption, it suffices, by two applications of GEN, to show:

$$(\varphi^{\beta}_{\alpha} \,\&\, \varphi^{\gamma}_{\alpha}) \rightarrow \beta = \gamma$$

So assume both φ^{β}_{α} and $\varphi^{\gamma}_{\alpha}$. From these assumptions, it follows from (ϑ) by (93.3) and (93.2), respectively, that $\beta = \tau$ and $\gamma = \tau$. We know $\beta\downarrow$ and $\gamma\downarrow$ by (39.2), and we also know $\tau\downarrow$ by hypothesis and (39.2), since τ is an arbitrarily chosen, primitive constant and thus not introduced by a definition. (The significance of these terms is also implied by the identities just established, by (107.1) and (107.2).) So, by the symmetry of identity, $\gamma = \tau$ implies $\tau = \gamma$ and, by the transitivity of identity, it follows that $\beta = \gamma$. ⋈

(129) Assume:

(a) $\forall\alpha(\varphi \rightarrow \Box\varphi)$

(b) $\exists!\alpha\varphi$

Now pick β to be a variable that doesn't occur free, and is substitutable for α, in φ. Then from (b), it follows by definition (127.1) of the uniqueness quantifier and our Rule $\equiv_{df}$E of Definiendum Elimination (90.2) that:

(ϑ) $\exists\alpha(\varphi \,\&\, \forall\beta(\varphi^{\beta}_{\alpha} \rightarrow \beta = \alpha))$

So suppose τ is an arbitrary such entity, i.e., suppose:

(ξ) $\varphi^{\tau}_{\alpha} \,\&\, \forall\beta(\varphi^{\beta}_{\alpha} \rightarrow \beta = \tau)$,

where τ is an arbitrarily chosen constant of the same type as α and β. Now to show $\exists!\alpha\Box\varphi$, we have to show, by definition (127.1) and Rule $\equiv_{df}$I of Definiendum Introduction (90.3), that $\exists\alpha(\Box\varphi \,\&\, \forall\beta(\Box\varphi^{\beta}_{\alpha} \rightarrow \beta = \alpha))$. By $\exists$I, it suffices to show that τ is a witness to this claim, i.e., show $\Box\varphi^{\tau}_{\alpha} \,\&\, \forall\beta(\Box\varphi^{\beta}_{\alpha} \rightarrow \beta = \tau)$. To show the first conjunct, note that since τ is a constant, we know both that $\tau\downarrow$ and that τ is substitutable for α in the matrix of (a). Hence, by Rule $\forall$E and MP, $\Box\varphi^{\tau}_{\alpha}$ is jointly implied by (a) and the first conjunct of (ξ). To show the second conjunct, it suffices by GEN to show $\Box\varphi^{\beta}_{\alpha} \rightarrow \beta = \tau$. So assume $\Box\varphi^{\beta}_{\alpha}$. Then, by the T schema (45.2), we have φ^{β}_{α}. But from this and the second conjunct of (ξ), it follows that $\beta = \tau$. Thus, we have established $\exists\alpha(\Box\varphi \,\&\, \forall\beta(\Box\varphi^{\beta}_{\alpha} \rightarrow \beta = \alpha))$. This conclusion remains once we discharge (ξ) by $\exists$E. ⋈

(130.1)★ For conditional proof, assume φ. Now, for reductio, assume $\neg\mathcal{A}\varphi$. Then by the right-to-left direction of the axiom governing negation and actuality, (44.1), $\mathcal{A}\neg\varphi$. Since it is axiomatic that $\mathcal{A}\psi \rightarrow \psi$ (43)★, it follows that $\neg\varphi$. Contradiction. So $\mathcal{A}\varphi$, by reductio, and $\varphi \rightarrow \mathcal{A}\varphi$, by conditional proof. ⋈

(130.2)★ Exercise.

(131) Assume, for conditional proof, $\mathcal{A}(\varphi \rightarrow \psi)$. From this and axiom (44.2), it follows that $\mathcal{A}\varphi \rightarrow \mathcal{A}\psi$, by biconditional syllogism. ⋈

(132) The T schema $\Box\varphi \rightarrow \varphi$ (45.2) is an axiom, and so are its closures. Hence, $\mathcal{A}(\Box\varphi \rightarrow \varphi)$ is an axiom. So by the previous theorem (131), $\mathcal{A}\Box\varphi \rightarrow \mathcal{A}\varphi$. However, as an instance of axiom (46.2), we know $\Box\varphi \equiv \mathcal{A}\Box\varphi$, from which it follows, by the definition of $\equiv$ (18.3) and &E (86.2.a), that $\Box\varphi \rightarrow \mathcal{A}\Box\varphi$. So by hypothetical syllogism, $\Box\varphi \rightarrow \mathcal{A}\varphi$. ⋈

(133.1) By applying the definition of $\equiv$ (18.3) to axiom (44.4) and then detaching the second conjunct by &E (86.2.b), we know:

(ζ) $\mathcal{A}\mathcal{A}\varphi \rightarrow \mathcal{A}\varphi$

Independently, by analogous reasoning, we may infer from axiom (44.2) that its right-to-left direction holds, but let us express this direction with different Greek metavariables so as to avoid clash of variables later:

$$(\mathcal{A}\chi \rightarrow \mathcal{A}\theta) \rightarrow \mathcal{A}(\chi \rightarrow \theta)$$

Now consider the instance of the above schema in which we let χ be the formula $\mathcal{A}\varphi$ and let θ be the formula φ:

(ξ) $(\mathcal{A}\mathcal{A}\varphi \rightarrow \mathcal{A}\varphi) \rightarrow \mathcal{A}(\mathcal{A}\varphi \rightarrow \varphi)$

It then follows from (ζ) and (ξ) that $\mathcal{A}(\mathcal{A}\varphi \rightarrow \varphi)$. ⋈

(133.2) By applying the definition of $\equiv$ (18.3) to axiom (44.4) and then detaching the first conjunct by &E (86.2.a), we know:

(ζ) $\mathcal{A}\varphi \rightarrow \mathcal{A}\mathcal{A}\varphi$

As in the previous theorem, we independently know the right-to-left direction of axiom (44.2) holds and we again express this with different Greek metavariables, to avoid clash of variables later:

$$(\mathcal{A}\chi \rightarrow \mathcal{A}\theta) \rightarrow \mathcal{A}(\chi \rightarrow \theta)$$

Now consider the following instance of the above schema, in which we let χ be the formula φ and let θ be the formula $\mathcal{A}\varphi$:

(ξ) $(\mathcal{A}\varphi \rightarrow \mathcal{A}\mathcal{A}\varphi) \rightarrow \mathcal{A}(\varphi \rightarrow \mathcal{A}\varphi)$

It then follows from (ζ) and (ξ) that $\mathcal{A}(\varphi \rightarrow \mathcal{A}\varphi)$. ⋈

(133.3) The principle of Adjunction (85.5) is $\varphi \rightarrow (\psi \rightarrow (\varphi \,\&\, \psi))$. Since this is a $\Box$-theorem, we may apply RN to obtain:

(c) $\Box(\varphi \rightarrow (\psi \rightarrow (\varphi \,\&\, \psi)))$

Now theorem (132) is that $\Box\chi \rightarrow \mathcal{A}\chi$, so it follows from (c) that:

(d) $\mathcal{A}(\varphi \rightarrow (\psi \rightarrow (\varphi \,\&\, \psi)))$

Then by theorem (131), (d) implies:

(e) $\mathcal{A}\varphi \rightarrow \mathcal{A}(\psi \rightarrow (\varphi \,\&\, \psi))$

Now if we consider only the consequent of (e), then we know, independently and by the same distribution law, that it obeys the principle:

(f) $\mathcal{A}(\psi \rightarrow (\varphi \,\&\, \psi)) \rightarrow (\mathcal{A}\psi \rightarrow \mathcal{A}(\varphi \,\&\, \psi))$

It now follows from (e) and (f) by hypothetical syllogism that:

(g) $\mathcal{A}\varphi \rightarrow (\mathcal{A}\psi \rightarrow \mathcal{A}(\varphi \,\&\, \psi))$

From (g), it follows by Importation (88.7.b) that $(\mathcal{A}\varphi \,\&\, \mathcal{A}\psi) \rightarrow \mathcal{A}(\varphi \,\&\, \psi)$. ⋈

(133.4) To avoid clash of variables, we may rewrite (133.3) as:

$$(\mathcal{A}\chi \,\&\, \mathcal{A}\theta) \to \mathcal{A}(\chi \,\&\, \theta)$$

Now consider the following instance of the above schema, in which we've set χ equal to $\mathcal{A}\varphi \to \varphi$ and set θ equal to $\varphi \to \mathcal{A}\varphi$:

$$(\vartheta)\ \ (\mathcal{A}(\mathcal{A}\varphi \to \varphi) \,\&\, \mathcal{A}(\varphi \to \mathcal{A}\varphi)) \to \mathcal{A}((\mathcal{A}\varphi \to \varphi) \,\&\, (\varphi \to \mathcal{A}\varphi))$$

But by &I, we may conjoin theorems (133.1) and (133.2) to produce:

$$(\xi)\ \ \mathcal{A}(\mathcal{A}\varphi \to \varphi) \,\&\, \mathcal{A}(\varphi \to \mathcal{A}\varphi)$$

Hence it follows from (ϑ) and (ξ) that:

$$(\zeta)\ \ \mathcal{A}((\mathcal{A}\varphi \to \varphi) \,\&\, (\varphi \to \mathcal{A}\varphi))$$

Note that we can't, at present, *conclude* that $\mathcal{A}(\mathcal{A}\varphi \equiv \varphi)$ from this last result *by definition*, since our rules for reasoning with definitions-by-$\equiv$ don't yet allow us to substitute the defined formula $\mathcal{A}\varphi \equiv \varphi$ for its definiens $(\mathcal{A}\varphi \to \varphi)\&(\varphi \to \mathcal{A}\varphi)$ when the latter occurs *within* a formula as a proper subformula.[433] But we can reason as follows. We know, as an instance of definition (18.3), that:

$$(\mathcal{A}\varphi \equiv \varphi) \equiv_{df} ((\mathcal{A}\varphi \to \varphi) \,\&\, (\varphi \to \mathcal{A}\varphi))$$

So by the Rule of Definition by Equivalence (72), we know that *any closure* of the following is a necessary axiom:

$$((\mathcal{A}\varphi \to \varphi) \,\&\, (\varphi \to \mathcal{A}\varphi)) \to (\mathcal{A}\varphi \equiv \varphi)$$

So the following closure is a necessary axiom and hence a theorem:

$$\mathcal{A}(((\mathcal{A}\varphi \to \varphi) \,\&\, (\varphi \to \mathcal{A}\varphi)) \to (\mathcal{A}\varphi \equiv \varphi))$$

So by theorem (131), we may distribute $\mathcal{A}$ over the conditional to obtain:

$$\mathcal{A}((\mathcal{A}\varphi \to \varphi) \,\&\, (\varphi \to \mathcal{A}\varphi)) \to \mathcal{A}(\mathcal{A}\varphi \equiv \varphi)$$

But from (ζ) and this last result, it follows that:

$$\mathcal{A}(\mathcal{A}\varphi \equiv \varphi) \qquad \bowtie$$

(134.1) Theorem (133.1) is:

$$(\vartheta)\ \ \mathcal{A}(\mathcal{A}\varphi \to \varphi)$$

But as an instance of the left-to-right direction of axiom (44.4), we know:

[433] In particular, the Rules of Definiendum Elimination (90.2) and Definiendum Introduction (90.3) don't permit the intersubstitution of definiens and definiendum when one or the other occurs as a proper subformula. Once we prove the special case (160.3) of the Rule of Substitution, however, we will be able to substitute definendum and definiens for one another whenever one occurs as a proper subformula within some formula.

(ξ) $\mathcal{A}(\mathcal{A}\varphi \to \varphi) \to \mathcal{A}\mathcal{A}(\mathcal{A}\varphi \to \varphi)$

So from (ϑ) and (ξ), it follows that:

(ζ) $\mathcal{A}\mathcal{A}(\mathcal{A}\varphi \to \varphi)$

If $\mathcal{A}\ldots\mathcal{A}(\mathcal{A}\varphi \to \varphi)$ is simply $\mathcal{A}\mathcal{A}(\mathcal{A}\varphi \to \varphi)$, then we're done. Otherwise, repeat the above of reasoning starting with (ζ) to conclude:

$\mathcal{A}\mathcal{A}\mathcal{A}(\mathcal{A}\varphi \to \varphi)$

And so on, as many times as needed to obtain the finite initial string of actuality operators in $\mathcal{A}\ldots\mathcal{A}(\mathcal{A}\varphi \equiv \varphi)$. ⋈

(134.2) – **(134.4)** (Exercises) [Note: The proofs are outlined in the text.]

(135) Suppose $\Gamma \vdash \varphi$, i.e., that there is a derivation of φ from Γ. We show by induction on the length of any such derivation that $\mathcal{A}\Gamma \vdash \mathcal{A}\varphi$, i.e., that there is a derivation of $\mathcal{A}\varphi$ from $\mathcal{A}\Gamma$.

Base Case. If $n = 1$, then the derivation of φ from Γ consists of a single formula, namely, φ itself. So, by the definition of $\Gamma \vdash \varphi$, φ must be in $\mathbf{\Lambda} \cup \Gamma$. So we have two cases: (A) φ is in $\mathbf{\Lambda}$ or (B) φ is in Γ.

> *Case A*: $\varphi \in \mathbf{\Lambda}$. Then φ is an axiom. So either (i) φ is a necessary axiom or (ii) φ is an instance of (43)★.[434] If (i), then the actualization of φ, i.e., $\mathcal{A}\varphi$, is an axiom, since all of the closures of necessary axioms are axioms. So $\vdash \mathcal{A}\varphi$ by (63.1) and $\mathcal{A}\Gamma \vdash \mathcal{A}\varphi$ by (63.3). If (ii), then as stipulated in (43)★, φ is either (a) $\mathcal{A}\psi \to \psi$, for some formula ψ, or (b) a universal closure of $\mathcal{A}\psi \to \psi$. (a) If φ is $\mathcal{A}\psi \to \psi$, then by (133.1), $\mathcal{A}(\mathcal{A}\psi \to \psi)$ is a theorem, i.e., $\vdash \mathcal{A}\varphi$. So $\mathcal{A}\Gamma \vdash \mathcal{A}\varphi$ by (63.3). (b) If φ is a universal closure of $\mathcal{A}\psi \to \psi$, i.e., if φ is a formula of the form $\forall\alpha_1\ldots\forall\alpha_n(\mathcal{A}\psi \to \psi)$, then by theorem (134.4), $\mathcal{A}\varphi$ is a theorem. So again, by (63.3), $\mathcal{A}\Gamma \vdash \mathcal{A}\varphi$.
>
> *Case B*: $\varphi \in \Gamma$. Then $\mathcal{A}\varphi$ is in $\mathcal{A}\Gamma$, by the definition of $\mathcal{A}\Gamma$. Hence by (63.2), it follows that $\mathcal{A}\Gamma \vdash \mathcal{A}\varphi$.

Inductive Case. Suppose that the derivation of φ from Γ is a sequence S of length n, where $n > 1$. Then either $\varphi \in \mathbf{\Lambda} \cup \Gamma$ or φ follows by MP from two previous members of the sequence, namely, $\psi \to \varphi$ and ψ. If $\varphi \in \mathbf{\Lambda} \cup \Gamma$, then using the reasoning in the base case, it follows that $\mathcal{A}\Gamma \vdash \mathcal{A}\varphi$. If φ follows from previous members $\psi \to \varphi$ and ψ by MP, then by the definition of a derivation, we know both that $\Gamma \vdash \psi \to \varphi$ and $\Gamma \vdash \psi$, where these are sequences of length less than n. Since our IH is that the theorem holds for all such derivations of length less than n, it follows that:

[434] Note that case (ii), which would apply to any other modally fragile axiom that may have been added to the system, should be omitted when the present proof is converted to a proof of the $\vdash_\Box$ form of Rule RA, which asserts that if $\Gamma \vdash_\Box \varphi$, then $\mathcal{A}\Gamma \vdash_\Box \mathcal{A}\varphi$.

(a) $\mathcal{A}\Gamma \vdash \mathcal{A}(\psi \rightarrow \varphi)$

(b) $\mathcal{A}\Gamma \vdash \mathcal{A}\psi$

Now since (44.2) is an axiom, we know:

$$\vdash \mathcal{A}(\psi \rightarrow \varphi) \equiv (\mathcal{A}\psi \rightarrow \mathcal{A}\varphi)$$

By definition (18.3), this is just:

$$\vdash (\mathcal{A}(\psi \rightarrow \varphi) \rightarrow (\mathcal{A}\psi \rightarrow \mathcal{A}\varphi)) \;\&\; ((\mathcal{A}\psi \rightarrow \mathcal{A}\varphi) \rightarrow \mathcal{A}(\psi \rightarrow \varphi))$$

So by &E (86.2.a), it follows that:

$$\vdash \mathcal{A}(\psi \rightarrow \varphi) \rightarrow (\mathcal{A}\psi \rightarrow \mathcal{A}\varphi)$$

It follows from this by (63.3) that:

$$\mathcal{A}\Gamma \vdash \mathcal{A}(\psi \rightarrow \varphi) \rightarrow (\mathcal{A}\psi \rightarrow \mathcal{A}\varphi)$$

So by (63.5), it follows from this and (a) that:

$$\mathcal{A}\Gamma \vdash \mathcal{A}\psi \rightarrow \mathcal{A}\varphi$$

And again by (63.5), it follows from this and (b) that:

$$\mathcal{A}\Gamma \vdash \mathcal{A}\varphi \qquad \bowtie$$

(137) Consider the set of axioms $\mathbf{\Lambda}$, which was defined in (59) as containing just the axioms asserted in Chapter 8. Then we prove:

Fact: If $\vdash \varphi$, then $\vdash_\Box \mathcal{A}\varphi$

Assume $\vdash \varphi$. We show, by induction on the length of any proof of φ that $\vdash_\Box \mathcal{A}\varphi$, i.e., that there is a modally strict proof of $\mathcal{A}\varphi$.

Base Case. $n = 1$. Then the proof of φ consists of a single formula, namely, φ itself. So by the definition of $\vdash \varphi$, φ must be an axiom in $\mathbf{\Lambda}$. So we have two cases: (*A*) φ is a necessary axiom (i.e., $\varphi \in \mathbf{\Lambda}_\Box$), or (*B*) φ is an instance of axiom (43)★ or a universal closure of such an instance.

> *Case A*: $\varphi \in \mathbf{\Lambda}_\Box$. Then by (63.1), it follows that $\vdash_\Box \varphi$. So by the $\vdash_\Box$ version of RA (135), which we didn't explicitly formulate given the convention in (67), it follows that $\vdash_\Box \mathcal{A}\varphi$.[435]

[435] Alternatively, from $\vdash_\Box \varphi$, it follows by (the official form of) RN (68) that $\vdash_\Box \Box\varphi$, which then implies $\vdash_\Box \mathcal{A}\varphi$, by the fact that (132), i.e., $\Box\varphi \rightarrow \mathcal{A}\varphi$, is a modally strict theorem.

Case B: φ is an instance of (43)$\star$ or a universal closure of such an instance. If φ is an instance of (43)$\star$, then φ has the form $\mathcal{A}\psi \to \psi$, for some formula ψ. Then by theorem (133.1), we know that $\mathcal{A}(\mathcal{A}\psi \to \psi)$ is a modally strict theorem, i.e., that $\vdash_\Box \mathcal{A}(\mathcal{A}\psi \to \psi)$, i.e., that $\vdash_\Box \mathcal{A}\varphi$. If φ is a universal closure of an instance of (43)$\star$, then by (134.4), we again know that $\vdash_\Box \mathcal{A}\varphi$.

Inductive case. If the proof of φ is a sequence of length n, where $n > 1$, then either $\varphi \in \mathbf{\Lambda}$ or φ follows from two previous members of the sequence, namely, ψ and $\psi \to \varphi$, by MP. If $\varphi \in \mathbf{\Lambda}$, then using the same reasoning as in the base case, $\vdash_\Box \mathcal{A}\varphi$. If φ follows from previous members $\psi \to \varphi$ and ψ, then by the definition of a proof, we know both $\vdash \psi$ and $\vdash \psi \to \varphi$. Then our Inductive Hypothesis implies:

(IH1) If $\vdash \psi$, then $\vdash_\Box \mathcal{A}\psi$

(IH2) If $\vdash \psi \to \varphi$, then $\vdash_\Box \mathcal{A}(\psi \to \varphi)$

Since we've established the antecedents of both IH1 and IH2, we know, respectively, that:

(ϑ) $\vdash_\Box \mathcal{A}\psi$

(ξ) $\vdash_\Box \mathcal{A}(\psi \to \varphi)$

But since (131) is a theorem, we know:

(ζ) $\vdash_\Box \mathcal{A}(\psi \to \varphi) \to (\mathcal{A}\psi \to \mathcal{A}\varphi)$

So it follows from (ξ) and (ζ) by (63.5) that $\vdash_\Box \mathcal{A}\psi \to \mathcal{A}\varphi$. And from this last fact and (ϑ), it follows by (63.5) that $\vdash_\Box \mathcal{A}\varphi$. ⋈

(138.1)$\star$ Theorem (130.2)$\star$ is $\mathcal{A}\varphi \equiv \varphi$. So it follows by a tautology for biconditionals (88.4.b) that $\neg\mathcal{A}\varphi \equiv \neg\varphi$. ⋈

(138.2)$\star$ If we substitute $\neg\varphi$ for φ on both sides of (138.1)$\star$, we obtain the following instance: $\neg\mathcal{A}\neg\varphi \equiv \neg\neg\varphi$. But it is a tautology that $\neg\neg\varphi \equiv \varphi$, by (88.3.b) and (88.2.e). So by a biconditional syllogism (89.3.e), it follows that $\neg\mathcal{A}\neg\varphi \equiv \varphi$. ⋈

(139.1) As an instance of the tautology (83), we know $\mathcal{A}\varphi \vee \neg\mathcal{A}\varphi$. So we reason by cases (86.4.a) from the two disjuncts. If $\mathcal{A}\varphi$, then by $\vee$I (86.3.a), $\mathcal{A}\varphi \vee \mathcal{A}\neg\varphi$. If $\neg\mathcal{A}\varphi$, then by axiom (44.1), it follows that $\mathcal{A}\neg\varphi$. So by $\vee$I (86.3.b), $\mathcal{A}\varphi \vee \mathcal{A}\neg\varphi$. Hence, $\mathcal{A}\varphi \vee \mathcal{A}\neg\varphi$. ⋈

(139.2) By theorem (133.3), &I and the definition of $\equiv$, it suffices to show just the left-to-right direction. ($\to$) A tautology of conjunction simplification (85.1) is $(\varphi \,\&\, \psi) \to \varphi$. By the Rule of Actualization (135), it follows that $\mathcal{A}((\varphi \,\&\, \psi) \to \varphi)$. So by theorem (131), we may infer:

(a) $\mathcal{A}(\varphi \,\&\, \psi) \to \mathcal{A}\varphi$

By analogous reasoning from the other tautology of conjunction simplification (85.2), i.e., $(\varphi \,\&\, \psi) \to \psi$, we may similarly infer:

(b) $\mathcal{A}(\varphi \,\&\, \psi) \to \mathcal{A}\psi$

Now for conditional proof, assume $\mathcal{A}(\varphi \,\&\, \psi)$. Then from (a) and (b), respectively, we may conclude both $\mathcal{A}\varphi$ and $\mathcal{A}\psi$. So by &I, $\mathcal{A}\varphi \,\&\, \mathcal{A}\psi$. ⋈

(139.3) As an instance of (139.2), we have:

(ϑ) $\mathcal{A}((\varphi \to \psi) \,\&\, (\psi \to \varphi)) \equiv (\mathcal{A}(\varphi \to \psi) \,\&\, \mathcal{A}(\psi \to \varphi))$

Note that we can't immediately infer the desired conclusion that $\mathcal{A}(\varphi \equiv \psi) \equiv (\mathcal{A}(\varphi \to \psi) \,\&\, \mathcal{A}(\psi \to \varphi))$, since we can't substitute a definiendum for a definiens when the latter occurs as a subformula. But we can reason as follows. Given definition (18.3) and the Rule of Definition by Equivalence (72), we know the following are theorems :

$$(\varphi \equiv \psi) \to ((\varphi \to \psi) \,\&\, (\psi \to \varphi))$$

$$((\varphi \to \psi) \,\&\, (\psi \to \varphi)) \to (\varphi \equiv \psi)$$

Hence by applying the Rule of Actualization to these theorems:

$$\mathcal{A}((\varphi \equiv \psi) \to ((\varphi \to \psi) \,\&\, (\psi \to \varphi)))$$

$$\mathcal{A}(((\varphi \to \psi) \,\&\, (\psi \to \varphi)) \to (\varphi \equiv \psi))$$

So by (131), we may infer from these two theorems, respectively:

$$\mathcal{A}(\varphi \equiv \psi) \to \mathcal{A}((\varphi \to \psi) \,\&\, (\psi \to \varphi))$$

$$\mathcal{A}((\varphi \to \psi) \,\&\, (\psi \to \varphi)) \to \mathcal{A}(\varphi \equiv \psi)$$

By combining these two results by &I and applying the definition of $\equiv$, it follows that:

$$\mathcal{A}(\varphi \equiv \psi) \equiv \mathcal{A}((\varphi \to \psi) \,\&\, (\psi \to \varphi))$$

From this last result and (ϑ), it follows by biconditional syllogism that:

$$\mathcal{A}(\varphi \equiv \psi) \equiv (\mathcal{A}(\varphi \to \psi) \,\&\, \mathcal{A}(\psi \to \varphi))$$ ⋈

(139.4) ($\to$) By theorem (131), we know:

(a) $\mathcal{A}(\varphi \to \psi) \to (\mathcal{A}\varphi \to \mathcal{A}\psi)$

(b) $\mathcal{A}(\psi \to \varphi) \to (\mathcal{A}\psi \to \mathcal{A}\varphi)$

So by Double Composition (88.8.d), we may conjoin the antecedents of (a) and (b) into a single conjunctive antecedent and conjoin the consequents of (a) and (b) into a single conjunctive consequent, to obtain:

(ϑ) $(\mathcal{A}(\varphi \to \psi) \,\&\, \mathcal{A}(\psi \to \varphi)) \to ((\mathcal{A}\varphi \to \mathcal{A}\psi) \,\&\, (\mathcal{A}\psi \to \mathcal{A}\varphi))$

By definition of $\equiv$, we independently know:

(ξ) $((\mathcal{A}\varphi \to \mathcal{A}\psi) \,\&\, (\mathcal{A}\psi \to \mathcal{A}\varphi)) \to (\mathcal{A}\varphi \equiv \mathcal{A}\psi)$

So by hypothetical syllogism from (ϑ) and (ξ), it follows that:

$$(\mathcal{A}(\varphi \to \psi) \,\&\, \mathcal{A}(\psi \to \varphi)) \to (\mathcal{A}\varphi \equiv \mathcal{A}\psi)$$

$(\leftarrow)$ Assume $\mathcal{A}\varphi \equiv \mathcal{A}\psi$, for conditional proof. Then, by definition of $\equiv$ and &E, it follows that $\mathcal{A}\varphi \to \mathcal{A}\psi$ and $\mathcal{A}\psi \to \mathcal{A}\varphi$. But by the right-to-left direction of axiom (44.2), the first implies $\mathcal{A}(\varphi \to \psi)$ and the second implies $\mathcal{A}(\psi \to \varphi)$. So by &I, we are done. ⋈

(139.5) $(\to)$ By biconditional syllogism (89.3.e) from theorems (139.3) and (139.4). ⋈

(139.6) $(\to)$ This direction is just axiom (46.1). $(\leftarrow)$ This direction is an instance of the T schema. ⋈

(139.7) Assume $\mathcal{A}\Box\varphi$. Then by the right-to-left direction of axiom (46.2), it follows that $\Box\varphi$. So by (132), it follows that $\mathcal{A}\varphi$. But then by (46.1), it follows that $\Box\mathcal{A}\varphi$. ⋈

(139.8) Assume $\Box\varphi$. From this and axiom (46.2), it follows by biconditional syllogism that $\mathcal{A}\Box\varphi$. From this latter and theorem (139.7), it follows that $\Box\mathcal{A}\varphi$. ⋈

(139.9) $(\to)$ Assume $\mathcal{A}(\varphi \vee \psi)$, for conditional proof. But now assume, for reductio, $\neg(\mathcal{A}\varphi \vee \mathcal{A}\psi)$. Then by a De Morgan's Law (88.5.d), it follows that $\neg\mathcal{A}\varphi \,\&\, \neg\mathcal{A}\psi$. By &E, we have both $\neg\mathcal{A}\varphi$ and $\neg\mathcal{A}\psi$. These imply, respectively, by axiom (44.1), that $\mathcal{A}\neg\varphi$ and $\mathcal{A}\neg\psi$. We may conjoin these by &I to produce $\mathcal{A}\neg\varphi \,\&\, \mathcal{A}\neg\psi$, and by an appropriate instance of theorem (139.2), namely, $\mathcal{A}(\neg\varphi \,\&\, \neg\psi) \equiv \mathcal{A}\neg\varphi \,\&\, \mathcal{A}\neg\psi$, it follows by biconditional syllogism that:

(a) $\mathcal{A}(\neg\varphi \,\&\, \neg\psi)$

Now, independently, by the commutativity of $\equiv$ (88.2.e), we may transform an instance of De Morgan's law (88.5.d) to obtain $(\neg\varphi \,\&\, \neg\psi) \equiv \neg(\varphi \vee \psi)$ as a theorem. So we may apply the Rule of Actualization to this instance to obtain:

(b) $\mathcal{A}((\neg\varphi \,\&\, \neg\psi) \equiv \neg(\varphi \vee \psi))$

Hence, from (b) it follows by an appropriate instance of (139.5) that:

(c) $\mathcal{A}(\neg\varphi \,\&\, \neg\psi) \equiv \mathcal{A}\neg(\varphi \vee \psi))$

From (a) and (c), it follows by biconditional syllogism that $\mathcal{A}\neg(\varphi \vee \psi)$. But by axiom (44.1), it follows that $\neg\mathcal{A}(\varphi \vee \psi)$, which contradicts our initial assumption. Hence, we may conclude by reductio (RAA) version (87.3) that $\mathcal{A}\varphi \vee \mathcal{A}\psi$. ($\leftarrow$) Exercise. ⋈

(139.10) We first prove some preliminary lemmas:

(ϑ) $\neg\mathcal{A}\forall\alpha\neg\varphi \equiv \neg\forall\alpha\mathcal{A}\neg\varphi$

> *Proof*. As an instance of axiom (44.3), we know $\mathcal{A}\forall\alpha\neg\varphi \equiv \forall\alpha\mathcal{A}\neg\varphi$. So by the tautology (88.4.b), $\neg\mathcal{A}\forall\alpha\neg\varphi \equiv \neg\forall\alpha\mathcal{A}\neg\varphi$.

(ξ) $\neg\forall\alpha\mathcal{A}\neg\varphi \equiv \neg\forall\alpha\neg\mathcal{A}\varphi$

> *Proof*. We know $\forall\alpha(\mathcal{A}\neg\varphi \equiv \neg\mathcal{A}\varphi)$, since it is a closure of axiom (44.1). So by a theorem of quantification theory (99.3), it follows that $\forall\alpha\mathcal{A}\neg\varphi \equiv \forall\alpha\neg\mathcal{A}\varphi$. Hence by our tautology (88.4.b), $\neg\forall\alpha\mathcal{A}\neg\varphi \equiv \neg\forall\alpha\neg\mathcal{A}\varphi$.

With these two lemmas, we may reason as follows. By Rule $\equiv$Df (90.1), the biconditional derived from definition (18.3) is a theorem. So by Rule RA, it follows that $\mathcal{A}(\exists\alpha\varphi \equiv \neg\forall\alpha\neg\varphi)$. By theorem (139.5), it follows that $\mathcal{A}\exists\alpha\varphi \equiv \mathcal{A}\neg\forall\alpha\neg\varphi$. We may therefore reason from this result by biconditional chaining as follows:

$$
\begin{array}{rcll}
\mathcal{A}\exists\alpha\varphi & \equiv & \mathcal{A}\neg\forall\alpha\neg\varphi & \\
& \equiv & \neg\mathcal{A}\forall\alpha\neg\varphi & \text{by axiom (44.1)} \\
& \equiv & \neg\forall\alpha\mathcal{A}\neg\varphi & \text{by } (\vartheta) \\
& \equiv & \neg\forall\alpha\neg\mathcal{A}\varphi & \text{by } (\xi) \\
& \equiv & \exists\alpha\mathcal{A}\varphi & \text{by definition of } \exists \qquad \bowtie
\end{array}
$$

(139.11) ($\rightarrow$) Assume $\mathcal{A}\forall\alpha(\varphi \equiv \psi)$. Then by axiom (44.3), it follows that $\forall\alpha\mathcal{A}(\varphi \equiv \psi)$. So by Rule $\forall$E (93.3), $\mathcal{A}(\varphi \equiv \psi)$. By the left-to-right direction of (139.5), we may infer $\mathcal{A}\varphi \equiv \mathcal{A}\psi$. But since we've reached this result from an assumption in which α doesn't occur free, we may apply GEN and conclude $\forall\alpha(\mathcal{A}\varphi \equiv \mathcal{A}\psi)$. ($\leftarrow$) Assume $\forall\alpha(\mathcal{A}\varphi \equiv \mathcal{A}\psi)$. Then by Rule $\forall$E (93.3), $\mathcal{A}\varphi \equiv \mathcal{A}\psi$. By the right-to-left direction of (139.5), $\mathcal{A}(\varphi \equiv \psi)$. But since we've reached this result from an assumption in which α doesn't occur free, we may apply GEN and conclude $\forall\alpha\mathcal{A}(\varphi \equiv \psi)$. Hence by axiom (44.3), $\mathcal{A}\forall\alpha(\varphi \equiv \psi)$. ⋈

(140)⋆ Let φ be any formula in which β is substitutable for α and doesn't occur free in φ. We prove both directions of the biconditional. ($\rightarrow$) Assume $\forall\beta(\mathcal{A}\varphi^{\beta}_{\alpha} \equiv \beta = \alpha)$. Now independently apply GEN to theorem (130.2)⋆ to obtain $\forall\alpha(\mathcal{A}\varphi \equiv \varphi)$, which by (99.11), implies $\forall\alpha(\varphi \equiv \mathcal{A}\varphi)$. Since β is substitutable for

α in φ and doesn't occur free in φ, it follows by (99.13), that $\forall\beta(\varphi \equiv \mathcal{A}\varphi)^{\beta}_{\alpha}$, i.e., $\forall\beta(\varphi^{\beta}_{\alpha} \equiv \mathcal{A}\varphi^{\beta}_{\alpha})$. And from this result and our assumption it follows by (99.10) that $\forall\beta(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha)$. ($\leftarrow$) By analogous reasoning. ⋈

(141)★ Consider any φ in which z is substitutable for x and doesn't occur free. By axiom (47), we know $x = \imath x\varphi \equiv \forall z(\mathcal{A}\varphi^{z}_{x} \equiv z = x)$. But by our previous theorem (140)★, we know that $\forall z(\mathcal{A}\varphi^{z}_{x} \equiv z = x) \equiv \forall z(\varphi^{z}_{x} \equiv z = x)$. Hence, by biconditional syllogism, it follows that $x = \imath x\varphi \equiv \forall z(\varphi^{z}_{x} \equiv z = x)$. ⋈

(142)★ Consider any φ in which z is substitutable for x and doesn't occur free. Then by the fundamental theorem (141)★ for descriptions, we know:

(ϑ) $x = \imath x\varphi \equiv \forall z(\varphi^{z}_{x} \equiv z = x)$

But, given our hypothesis that z is substitutable for x in φ and doesn't occur free in φ, we have as an instance of (126.3) that:

$$(\varphi \,\&\, \forall z(\varphi^{z}_{x} \rightarrow z = x)) \equiv \forall z(\varphi^{z}_{x} \equiv z = x)$$

From this last claim, it follows by the commutativity of $\equiv$ (88.2.e) that:

(ξ) $\forall z(\varphi^{z}_{x} \equiv z = x) \equiv (\varphi \,\&\, \forall z(\varphi^{z}_{x} \rightarrow z = x))$

So by biconditional syllogism (89.3.e) from (ϑ) and (ξ) it follows that:

$$x = \imath x\varphi \equiv (\varphi \,\&\, \forall z(\varphi^{z}_{x} \rightarrow z = x))$$ ⋈

(143)★ By hypothesis, (i) ψ is either an exemplification formula $\Pi^n\kappa_1 \dots \kappa_n$ ($n \geq 1$) or an encoding formula $\kappa_1 \dots \kappa_n\Pi^n$, (ii) x occurs in ψ and only as one or more of the κ_i ($1 \leq i \leq n$), and (iii) z is substitutable for x in φ and doesn't occur free in φ. We want to show:

(a) $\psi^{\imath x\varphi}_{x} \equiv \exists x(\varphi \,\&\, \forall z(\varphi^{z}_{x} \rightarrow z = x) \,\&\, \psi)$

Our strategy will be to use Hintikka's schema (142)★. Since z is substitutable for x in φ and doesn't occur free in φ, we know that Hintikka's schema applies to φ, so that we have:

$$x = \imath x\varphi \equiv (\varphi \,\&\, \forall z(\varphi^{z}_{x} \rightarrow z = x))$$

By GEN, it follows that:

(b) $\forall x(x = \imath x\varphi \equiv (\varphi \,\&\, \forall z(\varphi^{z}_{x} \rightarrow z = x)))$

(b) will be used in proving both directions of (a).

($\rightarrow$) Assume:

(c) $\psi^{\imath x\varphi}_{x}$

for conditional proof. Since ψ is, by hypothesis, an exemplification or encoding formula, it follows from (c) by axiom (39.5.a) and (39.5.b) that $\imath x\varphi\downarrow$. So, where y is some variable that doesn't occur free in $\imath x\varphi$, it follows by (121.1) that:

(d) $\exists y(y = \imath x\varphi)$

Assume that a is an arbitrary such object, so that we know $a = \imath x\varphi$. If we instantiate (b) to a by Rule $\forall$E (93.2), we obtain:

$$a = \imath x\varphi \equiv \varphi^a_x \ \& \ \forall z(\varphi^z_x \to z = a)$$

So by biconditional syllogism, it follows that:

(e) $\varphi^a_x \ \& \ \forall z(\varphi^z_x \to z = a)$

Note independently that we've established the symmetry of identity for objects (117.2), so that by GEN, we know $\forall x\forall y(x = y \to y = x)$. In this universal claim, we may instantiate $\forall x$ to a and, given that we already know $\imath x\varphi\downarrow$, instantiate $\forall y$ to $\imath x\varphi$, thereby inferring from the assumption that $a = \imath x\varphi$ that $\imath x\varphi = a$. From this and (c) it follows by Rule =E (110) that ψ^a_x (i.e., the result of substituting a for *all* the occurrences of $\imath x\varphi$ in $\psi^{\imath x\varphi}_x$). Conjoining this last result with (e) by &I we obtain:

$$\varphi^a_x \ \& \ \forall z(\varphi^z_x \to z = a) \ \& \ \psi^a_x$$

Hence, by $\exists$I, our desired conclusion follows:

$$\exists x(\varphi \ \& \ \forall z(\varphi^z_x \to z = x) \ \& \ \psi)$$

Since we've inferred this conclusion from the assumption that $a = \imath x\varphi$, where a is arbitrary, we may discharge the assumption to reach our conclusion from (d), by $\exists$E (102).

($\leftarrow$) Assume, for conditional proof:

(f) $\exists x(\varphi \ \& \ \forall z(\varphi^z_x \to z = x) \ \& \ \psi)$

Assume b is an arbitrary such object, so that we know:

(g) $\varphi^b_x \ \& \ \forall z(\varphi^z_x \to z = b) \ \& \ \psi^b_x$

By instantiating b into (b), we have:

(h) $b = \imath x\varphi \equiv \varphi^b_x \ \& \ \forall z(\varphi^z_x \to z = b)$

If we now detach the first two conjuncts of (g) from the third conjunct by one application of &E, we have both:

(i) $\varphi^b_x \ \& \ \forall z(\varphi^z_x \to z = b)$

(j) ψ^b_x

From (i) and (h) it follows by biconditional syllogism that:

(k) $b = \iota x\varphi$

From (k) and (j), it follows by Rule =E (110) that $\psi^{\iota x\varphi}_x$. Thus, we may discharge (g) to reach this same conclusion from (f) by $\exists$E (102). ⋈

(144.1)★ ($\rightarrow$) Assume $\iota x\varphi\downarrow$. Then by the left-to-right direction of (121.1), $\exists y(y = \iota x\varphi)$, where y is some variable that doesn't occur free in φ. Suppose a is such an object, so that we know $a = \iota x\varphi$. Now, independently, if we apply GEN to Hintikka's schema (142)★ and instantiate the resulting universal claim to the constant a, we know, by Rule $\forall$E (93.2), that $a = \iota x\varphi \equiv \varphi^a_x \& \forall z(\varphi^z_x \rightarrow z = a)$, where z is some variable that is substitutable for x in φ and that doesn't occur free in φ. So it follows by biconditional syllogism that $\varphi^a_x \& \forall z(\varphi^z_x \rightarrow z = a)$. Hence, by $\exists$I, it follows that $\exists x(\varphi \& \forall z(\varphi^z_x \rightarrow z = x))$. Given our choice of z, it follows by definition of the uniqueness quantifier (127.1) that $\exists!x\varphi$. This last conclusion remains once we discharge our assumption about a by $\exists$E. ($\leftarrow$) Use analogous reasoning in the reverse direction. ⋈

(144.2)★ (Exercise)

(145.1)★ Assume $x = \iota x\varphi$. Now pick a variable, say z, that is substitutable for x in φ. Then by Hintikka's schema (142)★, it follows that $\varphi \& \forall z(\varphi^z_x \rightarrow z = x)$. *A fortiori*, φ. ⋈

(145.2)★ By hypothesis, z is substitutable for x in φ. If we apply GEN to the previous theorem (145.1)★, it is a theorem that:

(ϑ) $\forall x(x = \iota x\varphi \rightarrow \varphi)$

Since z is substitutable for x in φ, it follows from the definition of *substitutable for* (15) that z is substitutable for x in the formula $x = \iota x\varphi \rightarrow \varphi$. Since z is a variable and substitutable for x in the matrix of (ϑ), we may instantiate it into (ϑ), by Rule $\forall$E (93.2), to infer $(x = \iota x\varphi \rightarrow \varphi)^z_x$. It then follows from the definition of ψ^z_x (14) that $(x = \iota x\varphi \rightarrow \varphi)^z_x$ just is $z = \iota x\varphi \rightarrow \varphi^z_x$ (exercise).[436] ⋈

(145.3)★ By hypothesis, $\iota x\varphi$ is substitutable for x in φ. Assume, for conditional proof, $\iota x\varphi\downarrow$. (Recall our convention, mentioned at the end of (20), that $\iota x\varphi\downarrow$ abbreviates $(\iota x\varphi)\downarrow$. We henceforth omit mention of this convention in subsequent proofs.) Now we saw, in the proof of the previous theorem, that the following is a ★-theorem:

(ϑ) $\forall x(x = \iota x\varphi \rightarrow \varphi)$

[436] Note that $(\iota x\varphi)^z_x$ is just $\iota x\varphi$, since any free occurrences of x in φ become bound in $\iota x\varphi$.

Since $\imath x\varphi$ is, by hypothesis, substitutable for x in φ, it must be substitutable for x in $x = \imath x\varphi \to \varphi$.[437] Hence the conditions for applying Rule ∀E (93.1) obtain and we may conclude from (ϑ) that $(x = \imath x\varphi \to \varphi)_x^{\imath x\varphi}$, i.e., $\imath x\varphi = \imath x\varphi \to \varphi_x^{\imath x\varphi}$. But our assumption $\imath x\varphi\downarrow$ also implies, by Rule =I (118), that $\imath x\varphi = \imath x\varphi$. Hence, $\varphi_x^{\imath x\varphi}$. ⋈

(145.4)⋆ (Exercise)

(146) Let φ be any formula in which β is substitutable for α and doesn't occur free. Axiom (44.3) is $\mathcal{A}\varphi \equiv \mathcal{A}\mathcal{A}\varphi$ and by the commutativity of $\equiv$, $\mathcal{A}\mathcal{A}\varphi \equiv \mathcal{A}\varphi$ is a theorem. By GEN, it follows, respectively, that $\forall\alpha(\mathcal{A}\varphi \equiv \mathcal{A}\mathcal{A}\varphi)$ and $\forall\alpha(\mathcal{A}\mathcal{A}\varphi \equiv \mathcal{A}\varphi)$ are also theorems. Given our hypothesis about β, the following alphabetic variants of our last two theorems are also theorems, respectively, by (99.13): $\forall\beta[(\mathcal{A}\varphi \equiv \mathcal{A}\mathcal{A}\varphi)_\alpha^\beta]$ and $\forall\beta[(\mathcal{A}\mathcal{A}\varphi \equiv \mathcal{A}\varphi)_\alpha^\beta]$. By the definition of substitution (14), it follows from these last two theorems, respectively, that:

(ξ) $\forall\beta(\mathcal{A}\varphi_\alpha^\beta \equiv \mathcal{A}\mathcal{A}\varphi_\alpha^\beta)$

(ζ) $\forall\beta(\mathcal{A}\mathcal{A}\varphi_\alpha^\beta \equiv \mathcal{A}\varphi_\alpha^\beta)$

Using these two facts, we now prove both directions of our biconditional. ($\rightarrow$) Assume $\forall\beta(\mathcal{A}\varphi_\alpha^\beta \equiv \beta = \alpha)$. From ($\zeta$) and this assumption, it follows that $\forall\beta(\mathcal{A}\mathcal{A}\varphi_\alpha^\beta \equiv \beta = \alpha)$, by (99.10). ($\leftarrow$) Assume $\forall\beta(\mathcal{A}\mathcal{A}\varphi_\alpha^\beta \equiv \beta = \alpha)$. From ($\xi$) and this assumption, it follows that $\forall\beta(\mathcal{A}\varphi_\alpha^\beta \equiv \beta = \alpha)$, also by (99.10). ⋈

(147.1) We may reason as follows, where z is chosen to be some variable substitutable for x in φ:

$$\begin{array}{rcll} x = \imath x\varphi & \equiv & \forall z(\mathcal{A}\varphi_x^z \equiv z = x) & \text{by axiom (47)} \\ & \equiv & \forall z(\mathcal{A}\mathcal{A}\varphi_x^z \equiv z = x) & \text{by theorem (146)} \\ & \equiv & x = \imath x\mathcal{A}\varphi & \text{by axiom (47)} \end{array}$$ ⋈

(147.2) Assume $\imath x\varphi\downarrow$. Note independently that it follows from the previous theorem (147.1) by GEN that $\forall x(x = \imath x\varphi \ \equiv\ x = \imath x\mathcal{A}\varphi)$. Note also (exercise) that $\imath x\varphi$ is substitutable for x in the matrix of this last universal claim. Hence, by Rule ∀E (93.1), we may infer $\imath x\varphi = \imath x\varphi \ \equiv\ \imath x\varphi = \imath x\mathcal{A}\varphi$. Moreover, given our assumption, it follows by Rule =I (118) that $\imath x\varphi = \imath x\varphi$. Hence $\imath x\varphi = \imath x\mathcal{A}\varphi$. ⋈

(148) Suppose z is substitutable for x and doesn't occur free in φ. Then by the axiom governing descriptions (47) we know:

(ϑ) $x = \imath x\varphi \ \equiv\ \forall z(\mathcal{A}\varphi_x^z \equiv z = x)$

[437] If no free variable in $\imath x\varphi$ is captured when $\imath x\varphi$ is substituted for all the free occurrences of x in φ, then since the initial occurrence of x in $x = \imath x\varphi \to \varphi$ doesn't fall under the scope of any quantifiers, no free variable in $\imath x\varphi$ is captured when $\imath x\varphi$ is substituted for all the free occurrences of x in $x = \imath x\varphi \to \varphi$.

But, given our hypothesis that z is substitutable for x in φ and doesn't occur free in φ, we have as an instance of (126.3) that:

$$(\mathcal{A}\varphi \,\&\, \forall z(\mathcal{A}\varphi^z_x \rightarrow z = x)) \equiv \forall z(\mathcal{A}\varphi^z_x \equiv z = x)$$

From this last claim, it follows by the commutativity of $\equiv$ (88.2.e) that:

$$(\xi)\ \ \forall z(\mathcal{A}\varphi^z_x \equiv z = x) \equiv (\mathcal{A}\varphi \,\&\, \forall z(\mathcal{A}\varphi^z_x \rightarrow z = x))$$

So by biconditional syllogism (89.3.e) from (ϑ) and (ξ) it follows that:

$$x = \imath x\varphi \equiv (\mathcal{A}\varphi \,\&\, \forall z(\mathcal{A}\varphi^z_x \rightarrow z = x)) \qquad \bowtie$$

(149.1) Assume $\mathcal{A}\forall x(\varphi \equiv \psi)$. Then by the left-to-right direction of (139.11), we know:

$$(\vartheta)\ \ \forall x(\mathcal{A}\varphi \equiv \mathcal{A}\psi)$$

Now we want to show $\forall x(x = \imath x\varphi \equiv x = \imath x\psi)$. By GEN, it suffices to show $x = \imath x\varphi \equiv x = \imath x\psi$. ($\rightarrow$) Assume $x = \imath x\varphi$. Without loss of generality, pick some variable, say z, that (a) is substitutable for x in both φ and ψ and (b) doesn't occur free in φ and ψ. Then by the axiom governing descriptions (47), our assumption $x = \imath x\varphi$ implies:

$$(\xi)\ \ \forall z(\mathcal{A}\varphi^z_x \equiv z = x)$$

Now we want to show $x = \imath x\psi$. Again, by the axiom for descriptions and our choice of z, it suffices to show $\forall z(\mathcal{A}\psi^z_x \equiv z = x)$. And by GEN, it suffices to show $\mathcal{A}\psi^z_x \equiv z = x$. Now since z is, by hypothesis, substitutable for x in both φ and ψ, it is substitutable for x in $\mathcal{A}\varphi \equiv \mathcal{A}\psi$. That is, it is substitutable for x in the matrix of (ϑ). Hence, since z is a variable, it follows from (ϑ) by Rule $\forall$E (93.2) that:

$$(\mathcal{A}\varphi \equiv \mathcal{A}\psi)^z_x$$

But, by definition of χ^z_x (14), the above is just $\mathcal{A}\varphi^z_x \equiv \mathcal{A}\psi^z_x$, which by the commutativity of $\equiv$ is equivalent to:

$$(\zeta)\ \ \mathcal{A}\psi^z_x \equiv \mathcal{A}\varphi^z_x$$

Now by Rule $\forall$E (93.3), (ξ) implies $\mathcal{A}\varphi^z_x \equiv z = x$. But by biconditional syllogism from (ζ) and this last result, it follows that $\mathcal{A}\psi^z_x \equiv z = x$, which is all that it remained to show. ($\leftarrow$) By analogous reasoning. $\bowtie$

(149.2) Assume both $\imath x\varphi\downarrow$ and $\mathcal{A}\forall x(\varphi \equiv \psi)$. From our second assumption, it follows, by (149.1), that:

$$\forall x(x = \imath x\varphi \equiv x = \imath x\psi)$$

But from the assumption $\imath x\varphi\downarrow$ and the fact (exercise) that $\imath x\varphi$ is substitutable for x in the matrix of this last claim, it follows by Rule $\forall$E (93.1) that:

$$\imath x\varphi = \imath x\varphi \equiv \imath x\varphi = \imath x\psi$$

But it also follows from our first assumption that $\imath x\varphi = \imath x\varphi$, by Rule =I (118). Hence $\imath x\varphi = \imath x\psi$. ⋈

(149.3) Assume both $\imath x\varphi\downarrow$ and $\Box\forall x(\varphi \equiv \psi)$. From our second assumption, it follows from (132) that $\mathcal{A}\forall x(\varphi \equiv \psi)$. Hence $\imath x\varphi = \imath x\psi$, by (149.2). ⋈

(149.4) (Exercise)

(149.5) Assume $\imath x\varphi\downarrow$ and that the variable y doesn't occur free in φ. Note independently that as an instance of theorem (124.2), we know $\forall z\Box(z = z)$. Since $\imath x\varphi$ is substitutable for z in the matrix of this universal claim, we can apply Rule $\forall$E (93.1) and conclude $\Box\imath x\varphi = \imath x\varphi$. From this last fact, the assumption that $\imath x\varphi\downarrow$, and the fact that y doesn't occur free in φ, it follows by Rule $\exists$I (101.1) that $\exists y\Box(y = \imath x\varphi)$. ⋈

(150.1)★ Assume $\forall x(\varphi \equiv \psi)$. Then by the fact that $\chi \rightarrow \mathcal{A}\chi$ (130.1)★, we know $\mathcal{A}\forall x(\varphi \equiv \psi)$. So by (149.1), it follows that $\forall x(x = \imath x\varphi \equiv x = \imath x\psi)$. ⋈

(150.2)★ Assume both $\imath x\varphi\downarrow$ and $\forall x(\varphi \equiv \psi)$. Then by the fact that $\chi \rightarrow \mathcal{A}\chi$ (130.1)★, the second assumption implies $\mathcal{A}\forall x(\varphi \equiv \psi)$. From this and our first assumption, it follow by (149.2) that $\imath x\varphi = \imath x\psi$. ⋈

(151) For both directions, follow the proof of (143)★, but instead of appealing to (142)★, appeal to the modally-strict version of Hintikka's schema (148) and reason with respect to $\mathcal{A}\varphi$ instead of φ. ⋈

(152.1) ($\rightarrow$) Assume $\imath x\varphi\downarrow$. Then by the left-to-right direction of an appropriate instance of (121.1), $\exists y(y = \imath x\varphi)$, where y is some variable that doesn't occur free in φ. Suppose a is such an object, so that we know $a = \imath x\varphi$. Now, independently, if we apply GEN to the axiom governing descriptions (47) and instantiate the resulting universal claim to the constant a, we know, by Rule $\forall$E (93.2):

$$a = \imath x\varphi \equiv \forall z(\mathcal{A}\varphi^z_x \equiv z = a)$$

where z is some variable that is substitutable for x in φ and that doesn't occur free in φ. So it follows by biconditional syllogism that $\forall z(\mathcal{A}\varphi^z_x \equiv z = a)$. Hence, by $\exists$I, it follows that $\exists x\forall z(\mathcal{A}\varphi^z_x \equiv z = x)$. Given our choice of z, it follows by a fact about the uniqueness quantifier (127.2) that $\exists!x\mathcal{A}\varphi$. This last conclusion remains once we discharge our assumption about a by $\exists$E. ($\leftarrow$) Use analogous reasoning in the reverse direction. ⋈

(152.2) Assume $x = \imath x\varphi$. Then, for some variable, say z, that is substitutable for x in φ and that doesn't occur free in φ, we know, by the modally strict version of Hintikka's schema (148), that $\mathcal{A}\varphi \,\&\, \forall z(\mathcal{A}\varphi^z_x \rightarrow z = x)$. *A fortiori*, $\mathcal{A}\varphi$. ⋈

(152.3) Suppose z is substitutable for x in φ. By applying GEN to theorem (152.2), we know it is a theorem that:

(ϑ) $\forall x(x = \imath x\varphi \rightarrow \mathcal{A}\varphi)$

Since z is, by hypothesis, substitutable for x in φ, it is substitutable for x in the formula $x = \imath x\varphi \rightarrow \mathcal{A}\varphi$. Since z is a variable substitutable for x in the matrix of (ϑ), we may instantiate it into (ϑ), by Rule $\forall$E (93.2), to conclude $z = \imath x\varphi \rightarrow \mathcal{A}\varphi^z_x$. ⋈

(152.4) Assume $\imath x\varphi\downarrow$. Note independently that it follows by GEN from (152.2) that:

(ϑ) $\forall x(x = \imath x\varphi \rightarrow \mathcal{A}\varphi)$

Since $\imath x\varphi$ is, by hypothesis, substitutable for x in φ, it must be substitutable for x in $x = \imath x\varphi \rightarrow \mathcal{A}\varphi$. Hence the conditions for applying Rule $\forall$E (93.1) obtain and we may conclude from (ϑ) that $\imath x\varphi = \imath x\varphi \rightarrow \mathcal{A}\varphi^{\imath x\varphi}_x$. But our assumption $\imath x\varphi\downarrow$ also implies, by Rule =I (118), that $\imath x\varphi = \imath x\varphi$. Hence, $\mathcal{A}\varphi^{\imath x\varphi}_x$. ⋈

(152.5) Assume $\imath x\varphi = \imath x\psi$. By (107.1) and (107.2), respectively, it follows that $\imath x\varphi\downarrow$ and $\imath x\psi\downarrow$. So, instantiating these terms into (117.2), we therefore know $\imath x\psi = \imath x\varphi$. From this last result and (152.2), it follows that $\mathcal{A}\varphi$, and from our assumption and (152.2), it follows that $\mathcal{A}\psi$. Hence $\mathcal{A}\varphi \equiv \mathcal{A}\psi$. So by (139.5), $\mathcal{A}(\varphi \equiv \psi)$. Since x isn't free in our assumption, it follows by GEN that $\forall x\mathcal{A}(\varphi \equiv \psi)$. Hence by axiom (44.3), $\mathcal{A}\forall x(\varphi \equiv \psi)$. ⋈

(153.1) Assume $\exists!x\Box\varphi$, where y doesn't occur free in φ and is substitutable for x in φ. By Rule $\equiv_{df}$E of Definiendum Elimination (90.2), it follows from definition (127.1) that:

(ζ) $\exists x(\Box\varphi \,\&\, \forall z(\Box\varphi^z_x \rightarrow z = x))$

where z is some variable that is substitutable for x in φ and that doesn't occur free in φ. Suppose b is an arbitrary such object, i.e., that:

(ϑ) $\Box\varphi^b_x \,\&\, \forall z(\Box\varphi^z_x \rightarrow z = b)$

Now we want to show that $\forall y(y = \imath x\varphi \rightarrow \varphi^y_x)$. But it suffices by GEN to show $y = \imath x\varphi \rightarrow \varphi^y_x$. So assume $y = \imath x\varphi$. If we apply GEN to the modally strict version of Hintikka's Schema (148) and instantiate the result to the variable y, then $y = \imath x\varphi$ implies:

(ξ) $\mathcal{A}\varphi^y_x \,\&\, \forall z(\mathcal{A}\varphi^z_x \rightarrow z = y)$

If we now instantiate b into the second conjunct of (ξ), we obtain $\mathcal{A}\varphi^b_x \rightarrow b = y$. But the first conjunct of (ϑ) implies, by theorem (132), that $\mathcal{A}\varphi^b_x$. So $b = y$. But the first conjunct of (ϑ) also implies φ^b_x, by the T schema (45.2). Hence, by

Rule =E, φ^y_x. Since this is what remained to show that $\forall y(y = \imath x\varphi \rightarrow \varphi^y_x)$, we've derived the latter from (ϑ). Hence, by $\exists$E, we've derived $\forall y(y = \imath x\varphi \rightarrow \varphi^y_x)$ from (ξ) and, thus, from our initial assumption. ⋈

(153.2) Assume $\forall x(\varphi \rightarrow \Box\varphi)$ and $\exists!x\varphi$. Then by (129), it follows that $\exists!x\Box\varphi$. But then by (153.1), we may conclude $\forall y(y = \imath x\varphi \rightarrow \varphi^y_x)$. ⋈

(154) Assume $\imath\nu\varphi\downarrow$. Then by the Variant of Rule =I (118.1), it follows that $\imath\nu\varphi = \imath\nu\varphi$. By definition (16.4), (16.5) and the ensuing discussion in (16), we know that since $\imath\nu\varphi$ and $(\imath\nu\varphi)'$ are alphabetically-variant terms, then $\imath\nu\varphi = \imath\nu\varphi$ and $\imath\nu\varphi = (\imath\nu\varphi)'$ are alphabetically variant formulas. So it follows by the Rule of Alphabetic Variants (114) that $\imath\nu\varphi = (\imath\nu\varphi)'$. ⋈

(157.1) [We prove only the stronger form, since the weaker form follows trivially by (62.1).] Assume $\Gamma \vdash_\Box \varphi \rightarrow \psi$, i.e., that there is a modally-strict derivation of $\varphi \rightarrow \psi$ from Γ. Then the conditions of the strong form of RN are met and we may apply RN to conclude $\Box\Gamma \vdash_\Box \Box(\varphi \rightarrow \psi)$. Since instances of the K schema are necessary axioms, we know $\vdash_\Box \Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$, by (63.1). So $\Box\Gamma \vdash_\Box \Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$, by (63.3), and by (63.6), it follows that $\Box\Gamma \vdash_\Box \Box\varphi \rightarrow \Box\psi$. ⋈

(157.2) [We prove only the stronger form, since the weaker form follows trivially by (62.1).] Assume $\Gamma \vdash_\Box \varphi \rightarrow \psi$, i.e., that there is a modally-strict derivation of $\varphi \rightarrow \psi$ from Γ. Since the metarules of contraposition (80) apply generally to all derivations, they apply to modally-strict derivations, and so it follows by (80.1) that $\Gamma \vdash_\Box \neg\psi \rightarrow \neg\varphi$. Hence by Rule RM (157.1), it follows that $\Box\Gamma \vdash_\Box \Box\neg\psi \rightarrow \Box\neg\varphi$. So, again by our metarule of contraposition (80.1), it follows that:

(ϑ) $\Box\Gamma \vdash_\Box \neg\Box\neg\varphi \rightarrow \neg\Box\neg\psi$

Now by the definition of $\Diamond$ and the $\vdash_\Box$ form of the Rule of Definition by Equivalence (72), we independently know:[438]

$$\Box\Gamma \vdash_\Box \neg\Box\neg\psi \rightarrow \Diamond\psi$$

Hence it follows from this last result and (ϑ) by a corollary (76.1) to the Deduction Theorem that:

(ξ) $\Box\Gamma \vdash_\Box \neg\Box\neg\varphi \rightarrow \Diamond\psi$

Moreover, again by the definition of $\Diamond$ and Rule of Definition by Equivalence, we independently know:

$$\Box\Gamma \vdash_\Box \Diamond\varphi \rightarrow \neg\Box\neg\varphi$$

[438] Recall that though the Rule of Definition by Equivalence (72) was formulated for $\vdash$, it also holds for $\vdash_\Box$; we omitted the statement of the rule for $\vdash_\Box$, by convention (67).

From this last result and (ξ), it follows by the same corollary (76.1) to the Deduction Theorem that $\Box\Gamma \vdash_{\Box} \Diamond\varphi \rightarrow \Diamond\psi$. ⋈

(157.3) Assume $\Gamma \vdash_{\Box} \varphi \equiv \psi$. Then, by the definition of $\equiv$ and the $\vdash_{\Box}$ form of Rule $\equiv_{df}$E of Definiendum Elimination (90.2):

$$\Gamma \vdash_{\Box} (\varphi \rightarrow \psi) \,\&\, (\psi \rightarrow \varphi)$$

So by two applications of Rule &E, we have both:

$$\Gamma \vdash_{\Box} \varphi \rightarrow \psi$$

$$\Gamma \vdash_{\Box} \psi \rightarrow \varphi$$

By applying Rule RM (157.1)to each, we have:

$$\Gamma \vdash_{\Box} \Box\varphi \rightarrow \Box\psi$$

$$\Gamma \vdash_{\Box} \Box\psi \rightarrow \Box\varphi$$

Hence, by Rule &I:

$$\Gamma \vdash_{\Box} (\Box\varphi \rightarrow \Box\psi) \,\&\, (\Box\psi \rightarrow \Box\varphi)$$

So by the Rule of Definiendum Introduction (90.3):

$$\Box\Gamma \vdash_{\Box} \Box\varphi \equiv \Box\psi \qquad ⋈$$

(157.4) (Exercise)

(158.1) By the first axiom (38.1) governing conditionals, we have $\varphi \rightarrow (\psi \rightarrow \varphi)$. Since this is a $\Box$-theorem, we may apply RM (157.1) to conclude $\Box\varphi \rightarrow \Box(\psi \rightarrow \varphi)$. ⋈

(158.2) Since the tautology (77.3), i.e., $\neg\varphi \rightarrow (\varphi \rightarrow \psi)$, is a modally strict theorem, it follows by RM (157.1) that $\Box\neg\varphi \rightarrow \Box(\varphi \rightarrow \psi)$. ⋈

(158.3) ($\rightarrow$) A tautology of conjunction simplification (85.1) is $(\varphi \,\&\, \psi) \rightarrow \varphi$. Since this is a $\Box$-theorem, we may apply RM (157.1) to obtain:

(a) $\Box(\varphi \,\&\, \psi) \rightarrow \Box\varphi$

By analogous reasoning from $(\varphi \,\&\, \psi) \rightarrow \psi$ (85.2), we obtain:

(b) $\Box(\varphi \,\&\, \psi) \rightarrow \Box\psi$

Assume $\Box(\varphi \,\&\, \psi)$ for conditional proof. Then from (a) and (b), respectively, we may infer $\Box\varphi$ and $\Box\psi$. Hence by &I, $\Box\varphi \,\&\, \Box\psi$. ($\leftarrow$) The principle of Adjunction (85.5) is $\varphi \rightarrow (\psi \rightarrow (\varphi \,\&\, \psi))$. Since this is a $\Box$-theorem, we may apply RM (157.1) to obtain:

(c) $\Box\varphi \to \Box(\psi \to (\varphi \,\&\, \psi))$

The consequent of (c) can be used to form an instance of K (45.1):

(d) $\Box(\psi \to (\varphi \,\&\, \psi)) \to (\Box\psi \to \Box(\varphi \,\&\, \psi))$

By hypothetical syllogism (76.3), if follows from (c) and (d) that $\Box\varphi \to (\Box\psi \to \Box(\varphi \,\&\, \psi))$. Then by Importation (88.7.b), it follows that $(\Box\varphi \,\&\, \Box\psi) \to \Box(\varphi \,\&\, \psi)$. ⋈

(158.4) As an instance of (158.3), we have:

(ϑ) $\Box((\varphi \to \psi) \,\&\, (\psi \to \varphi)) \equiv (\Box(\varphi \to \psi) \,\&\, \Box(\psi \to \varphi))$

Independently, we know, by the $\vdash_\Box$ form of Rule $\equiv$Df (90.1) of Equivalence by Definition, that the following is a modally strict theorem, given the definition of $\equiv$:

$$(\varphi \equiv \psi) \equiv ((\varphi \to \psi) \,\&\, (\psi \to \varphi))$$

So by Rule RE (157.3):

(ξ) $\Box(\varphi \equiv \psi) \equiv \Box((\varphi \to \psi) \,\&\, (\psi \to \varphi))$

But then it follows from (ξ) and (ϑ) by biconditional syllogism that:

$$\Box(\varphi \equiv \psi) \equiv (\Box(\varphi \to \psi) \,\&\, \Box(\psi \to \varphi)) \qquad ⋈$$

(158.5) The following are both instances of the K axiom (45.1):

$$\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$$

$$\Box(\psi \to \varphi) \to (\Box\psi \to \Box\varphi)$$

So by Double Composition (88.8.d), we may conjoin the two antecedents into a single conjunctive antecedent and conjoin the two consequents into a single conjunctive consequent, to obtain:

$$(\Box(\varphi \to \psi) \,\&\, \Box(\psi \to \varphi)) \to ((\Box\varphi \to \Box\psi) \,\&\, (\Box\psi \to \Box\varphi))$$

But by the Rule of Definition by Equivalence (72) and the definition of $\equiv$, we know:

$$((\Box\varphi \to \Box\psi) \,\&\, (\Box\psi \to \Box\varphi)) \to (\Box\varphi \equiv \Box\psi)$$

So by hypothetical syllogism:

$$(\Box(\varphi \to \psi) \,\&\, \Box(\psi \to \varphi)) \to (\Box\varphi \equiv \Box\psi) \qquad ⋈$$

(158.6) Theorem (158.4) is $\Box(\varphi \equiv \psi) \equiv (\Box(\varphi \to \psi) \,\&\, \Box(\psi \to \varphi))$. From this, it follows, by the definition of $\equiv$ and the Rules $\equiv_{df}$E of Definiendum Elimination (90.2) and &E (86.2.a) that $\Box(\varphi \equiv \psi) \to (\Box(\varphi \to \psi) \,\&\, \Box(\psi \to \varphi))$. But from this and (158.5), it follows by hypothetical syllogism (76.3) that $\Box(\varphi \equiv \psi) \to (\Box\varphi \equiv \Box\psi)$. ⋈

(158.7) Assume $(\Box\varphi \,\&\, \Box\psi) \vee (\Box\neg\varphi \,\&\, \Box\neg\psi)$. Then we reason by cases (86.4.a).

Assume $\Box\varphi \,\&\, \Box\psi$. Then by &E, both $\Box\varphi$ and $\Box\psi$. From the latter, it follows by (158.1) that $\Box(\varphi \to \psi)$. From the former, it follows by (158.1) that $\Box(\psi \to \varphi)$. Hence, by &I, it follows that $\Box(\varphi \to \psi) \,\&\, \Box(\psi \to \varphi)$. This last conclusion and the right-to-left direction of (158.4) jointly imply $\Box(\varphi \equiv \psi)$, by biconditional syllogism.

Now assume $\Box\neg\varphi \,\&\, \Box\neg\psi$. By the right-to-left direction of (158.3), it follows that $\Box(\neg\varphi \,\&\, \neg\psi)$. Note independently that it is easy to establish, by conditional proof, $\vee$I, and the right-to-left direction of (88.4.g), that $(\neg\varphi \,\&\, \neg\psi) \to (\varphi \equiv \psi)$. Since this is a modally strict theorem, it follows by RM (157.1) that $\Box(\neg\varphi \,\&\, \neg\psi) \to \Box(\varphi \equiv \psi)$. Since we've already established the antecedent, $\Box(\varphi \equiv \psi)$ follows by Modus Ponens. ⋈

(158.8) Assume $\Box(\varphi \,\&\, \psi)$. From this and (158.3), it follows by $\equiv$E (89.3.a) that $\Box\varphi \,\&\, \Box\psi$. So by (158.7), $\Box(\varphi \equiv \psi)$. ⋈

(158.9) Assume $\Box(\neg\varphi \& \neg\psi)$. Independently, we leave it as an exercise to prove, using the right-to-left direction of the modally strict theorem (88.4.g) and $\vee$I, that $(\neg\varphi \,\&\, \neg\psi) \to (\varphi \equiv \psi)$ is a modally-strict theorem. Hence by Rule RM it follows that:

$$\Box(\neg\varphi \equiv \neg\psi) \to \Box(\varphi \equiv \psi)$$

From this and our assumption, it follows that $\Box(\varphi \equiv \psi)$. ⋈

(158.10) Since the tautology $\varphi \equiv \neg\neg\varphi$ (88.3.b) is a $\Box$-theorem, it follows by Rule RE that $\Box\varphi \equiv \Box\neg\neg\varphi$. ⋈

(158.11) ($\to$) Assume $\neg\Box\varphi$, for conditional proof. We want to show $\Diamond\neg\varphi$. By definition of $\Diamond$ and Rule $\equiv_{df}$I, we have to show $\neg\Box\neg\neg\varphi$. For reductio, assume $\Box\neg\neg\varphi$. From this and (158.10), it follows by biconditional syllogism that $\Box\varphi$, which contradicts our initial assumption. ($\leftarrow$) Assume $\Diamond\neg\varphi$. Then by definition of $\Diamond$ and Rule $\equiv_{df}$I, $\neg\Box\neg\neg\varphi$. We want to show $\neg\Box\varphi$. So, for reductio, assume $\Box\varphi$. From this and (158.10), it follows by biconditional syllogism that $\Box\neg\neg\varphi$, which is a contradiction. ⋈

(158.12) ($\to$) Assume $\Box\varphi$. For reductio, assume $\Diamond\neg\varphi$. From this and (158.11), it follows by biconditional syllogism that $\neg\Box\varphi$, which contradicts our initial assumption. ($\leftarrow$) Assume $\neg\Diamond\neg\varphi$, for conditional proof. For reductio, assume

$\neg\Box\varphi$. From this and (158.11), it follows by biconditional syllogism that $\Diamond\neg\varphi$, which contradicts our initial assumption. ⋈

(158.13) By the special case of (63.2), we know $(\varphi \to \psi) \vdash_{\Box} (\varphi \to \psi)$. So by RM$\Diamond$ (157.2), it follows that $\Box(\varphi \to \psi) \vdash_{\Box} \Diamond\varphi \to \Diamond\psi$. Hence, by the $\vdash_{\Box}$ version of the Deduction Theorem (75), it is a modally strict theorem that $\Box(\varphi \to \psi) \to \Diamond\varphi \to \Diamond\psi$. ⋈

(158.14) By (158.11), it is a $\Box$-theorem that $\neg\Box\varphi \equiv \Diamond\neg\varphi$. Hence, by RE, it follows that $\Box\neg\Box\varphi \equiv \Box\Diamond\neg\varphi$. By the relevant instance of a biconditional tautology (88.4.b), we can negate both sides to obtain: $\neg\Box\neg\Box\varphi \equiv \neg\Box\Diamond\neg\varphi$. But by definition of $\Diamond$ and Rule $\equiv$Df (90.1), we know $\Diamond\Box\varphi \equiv \neg\Box\neg\Box\varphi$. So by biconditional syllogism, it follows that $\Diamond\Box\varphi \equiv \neg\Box\Diamond\neg\varphi$. ⋈

(158.15) By simple conditional proofs and the rules for $\vee$I (86.3.a) and (86.3.b), we can establish the following $\Box$-theorems:

$$\varphi \to (\varphi \vee \psi)$$

$$\psi \to (\varphi \vee \psi)$$

Hence it follows by Rule RM (157.1) that:

$$\Box\varphi \to \Box(\varphi \vee \psi)$$

$$\Box\psi \to \Box(\varphi \vee \psi)$$

So by an appropriate instance of the tautology (88.8.c), it follows that $(\Box\varphi \vee \Box\psi) \to \Box(\varphi \vee \psi)$. ⋈

(158.16) Assume:

(ϑ) $\Box\varphi \,\&\, \Diamond\psi$

Now independently, by (85.5), we know that $\varphi \to (\psi \to (\varphi \,\&\, \psi)$ is a modally strict theorem. So by Rule RM, $\Box\varphi \to \Box(\psi \to (\varphi \,\&\, \psi))$. So from this last result and the first conjunct of (ϑ), $\Box(\psi \to (\varphi \,\&\, \psi))$. So by the K$\Diamond$ schema (158.13), $\Diamond\psi \to \Diamond(\varphi \,\&\, \psi)$. This result and the second conjunct of (ϑ) imply $\Diamond(\varphi \,\&\, \psi)$. ⋈

(159.1) Our global assumption is:

(ξ) $\vdash \Box(\psi \equiv \chi)$

Since instances of the T schema (45.2) are axioms, we know:

$$\vdash \Box(\psi \equiv \chi) \to (\psi \equiv \chi)$$

by (63.1). From this and our global assumption, it follows by (63.6) that:

(ϑ) $\vdash \psi \equiv \chi$

We frequently appeal to either (ξ) or (ϑ) in establishing the following cases of the consequent of the rule:

(.a) Show $\vdash \neg\psi \equiv \neg\chi$. Since instances of (88.4.b) are theorems, we know:

$$\vdash \psi \equiv \chi \equiv \neg\psi \equiv \neg\chi$$

From this and (ϑ), it follows by biconditional syllogism (89.3.a) that:

$$\vdash \neg\psi \equiv \neg\chi$$

(.b) Show $\vdash (\psi \to \theta) \equiv (\chi \to \theta)$. Note that since instances of (88.4.c) are theorems, we know:

$$\vdash (\psi \equiv \chi) \to ((\psi \to \theta) \equiv (\chi \to \theta))$$

From this and (ϑ), it follows by (63.6) that $\vdash (\psi \to \theta) \equiv (\chi \to \theta)$.

(.c) Show $\vdash (\theta \to \psi) \equiv (\theta \to \chi)$. By reasoning analogous to the previous case, but starting with an instance of (88.4.d) instead of (88.4.c).

(.d) Show $\vdash \forall\alpha\psi \equiv \forall\alpha\chi$. From ($\vartheta$), it follows by GEN that $\vdash \forall\alpha(\psi \equiv \chi)$. But by our proof of (99.3), we know:

$$\vdash \forall\alpha(\psi \equiv \chi) \to (\forall\alpha\psi \equiv \forall\alpha\chi)$$

Hence it follows by (63.6) that $\vdash \forall\alpha\psi \equiv \forall\alpha\chi$.

(.e) Show $\vdash [\lambda\, \psi] \equiv [\lambda\, \chi]$. Since the instances of (111.6) are theorems, we know:

$$\vdash (\psi \equiv \chi) \equiv ([\lambda\, \psi] \equiv [\lambda\, \chi])$$

From this fact and (ϑ), it follows by biconditional syllogism (89.3.a) that $\vdash [\lambda\, \psi] \equiv [\lambda\, \chi]$.

(.f) Show $\vdash \mathcal{A}\psi \equiv \mathcal{A}\chi$. Our proof of (132) establishes that:

$$\vdash \Box(\psi \equiv \chi) \to \mathcal{A}(\psi \equiv \chi)$$

From this and our global assumption (ξ), it follows that:

(ζ) $\vdash \mathcal{A}(\psi \equiv \chi)$, by (63.6).[439] But by (139.5), we know that $\mathcal{A}(\psi \equiv \chi) \equiv (\mathcal{A}\psi \equiv \mathcal{A}\chi)$, so that we know:[440]

$$\vdash \mathcal{A}(\psi \equiv \chi) \equiv (\mathcal{A}\psi \equiv \mathcal{A}\chi)$$

Hence it follows that $\vdash \mathcal{A}\psi \equiv \mathcal{A}\chi$, by (89.3.a).

(.g) Show $\vdash \Box\psi \equiv \Box\chi$. By our proof of (158.6), we know:

$$\vdash \Box(\psi \equiv \chi) \rightarrow (\Box\psi \equiv \Box\chi)$$

But it follows from this and our global assumption (ξ) that $\vdash \Box\psi \equiv \Box\chi$, by biconditional syllogism (89.3.a). ⋈

(159.2) Assume:

(ϑ) $\vdash \Box(\psi \equiv \chi)$

and let φ' be the result of substituting the formula χ for zero or more occurrences of ψ where the latter is a subformula of φ. We then show, by induction on the complexity of φ, that $\vdash \varphi \equiv \varphi'$.

However, we may put aside the following subcases that occur in the base case and the inductive cases:

(ζ) If no occurrences of ψ in φ are replaced by χ, then $\varphi' = \varphi$, and we simply have to show $\vdash \varphi \equiv \varphi$. But $\varphi \equiv \varphi$ is theorem (88.3.a).

(ξ) If ψ is a subformula of φ because $\psi = \varphi$, then $\varphi' = \chi$ and (ϑ) becomes $\vdash \Box(\varphi \equiv \varphi')$. Since instances of the T schema (45.2) are axioms, we know $\vdash \Box(\varphi \equiv \varphi') \rightarrow (\varphi \equiv \varphi')$, by (63.1). It follows that $\vdash \varphi \equiv \varphi'$, by (63.6).

Given (ζ) and (ξ), we need to reason only in those cases where (a) φ' is the result of substituting χ for at least one occurrence of ψ in φ, and (b) ψ is a proper subformula of φ (i.e., a subformula of φ not identical with φ).

Base Case. By our BNF (4), φ is (a) an exemplification formula of the form $\Pi^n\kappa_1\ldots\kappa_n$ $(n \geq 1)$ or (b) an encoding formula of the form $\kappa_1\ldots\kappa_n\Pi^n$ $(n \geq 1)$ or (c) a 0-ary relation constant or variable. In all three cases, φ has no proper subformulas and so ψ can't be a proper subformula of φ. So these cases are covered by (ζ) and (ξ) above.

[439] Although we could have established $\vdash \mathcal{A}(\psi \equiv \chi)$ by citing (ϑ) and appealing to theorem (130.2)★, we have refrained from doing so. If we had done so, this rule and the subsequently derivable the Rules of Substitution would have become non-strict metarules, since their justification would depend on an axiom that is modally fragile. Any conclusions drawn using the metarules derived in this manner would be ★-theorems. By appealing to (132), we prove the present case without an appeal to any ★-theorems.

[440] Again, (139.5) could have been proved using an appeal to (130.2)★, but for the reasons given in footnote 439, we are relying on the proof that makes no appeal to ★-theorems.

Inductive Case 1. $\varphi = [\lambda\, \theta]$. Then φ', i.e., $[\lambda\, \theta]'$, must be $[\lambda\, \theta']$. So our IH implies:

$$\text{If} \vdash \Box(\psi \equiv \chi), \text{then} \vdash \theta \equiv \theta'$$

It follows from this and (ϑ) that $\vdash \theta \equiv \theta'$. But from our proof of (111.6), it follows *a fortiori* that:

$$\vdash (\theta \equiv \theta') \rightarrow ([\lambda\, \theta] \equiv [\lambda\, \theta'])$$

Hence it follows by (63.6) that $\vdash [\lambda\, \theta] \equiv [\lambda\, \theta']$, i.e., $\varphi \equiv \varphi'$.

Inductive Case 2. $\varphi = \neg\theta$. Then φ', i.e., $(\neg\theta)'$ must be $\neg(\theta')$, which we henceforth write simply as $\neg\theta'$. Now our IH implies:

$$\text{If} \vdash \Box(\psi \equiv \chi), \text{then} \vdash \theta \equiv \theta'$$

So it follows from this and (ϑ) that $\vdash \theta \equiv \theta'$. Since we've proved instances of the tautology (88.4.b), we know:

$$\vdash (\theta \equiv \theta') \equiv (\neg\theta \equiv \neg\theta')$$

From this and $\vdash \theta \equiv \theta'$, it follows by biconditional syllogism (89.3.a) that $\vdash \neg\theta \equiv \neg\theta'$, i.e., $\vdash \varphi \equiv \varphi'$.

Inductive Case 3. $\varphi = \theta \rightarrow \omega$. Then φ', i.e., $(\theta \rightarrow \omega)$, must be $\theta' \rightarrow \omega'$. Our IHs are:

$$\text{If} \vdash \Box(\psi \equiv \chi), \text{then} \vdash \theta \equiv \theta'$$

$$\text{If} \vdash \Box(\psi \equiv \chi), \text{then} \vdash \omega \equiv \omega'$$

So it follows from these and (ϑ) that $\vdash \theta \equiv \theta'$ and $\vdash \omega \equiv \omega'$. Hence by &I (86.1), it follows that:

(a) $\vdash (\theta \equiv \theta') \,\&\, (\omega \equiv \omega)'$

But one can prove (as an exercise) the tautology:

$$((\theta \equiv \theta') \,\&\, (\omega \equiv \omega')) \rightarrow ((\theta \rightarrow \omega) \equiv (\theta' \rightarrow \omega'))$$

It therefore follows from the theoremhood of this tautology and (a), by (63.6), that:

$$\vdash (\theta \rightarrow \omega) \equiv (\theta' \rightarrow \omega')$$

i.e., $\vdash \varphi \equiv \varphi'$.

Inductive Case 4. $\varphi = \forall\alpha\theta$. Then φ', i.e., $(\forall\alpha\theta)'$, must be $\forall\alpha\theta'$. Our IH implies:

$$\text{If} \vdash \Box(\psi \equiv \chi), \text{then} \vdash \theta \equiv \theta'$$

So it follows from this and (ϑ) that $\vdash \theta \equiv \theta'$. So by GEN it follows that $\vdash \forall\alpha(\theta \equiv \theta')$. But by our proof of (99.3), we know:

$$\vdash \forall\alpha(\theta \equiv \theta') \rightarrow (\forall\alpha\theta \equiv \forall\alpha\theta')$$

Hence it follows by biconditional syllogism (89.3.a) that $\vdash \forall\alpha\theta \equiv \forall\alpha\theta'$, i.e., $\vdash \varphi \equiv \varphi'$.

Inductive Case 5. $\varphi = \mathcal{A}\theta$. Then φ', i.e., $(\mathcal{A}\theta)'$ must be $\mathcal{A}\theta'$. Our IH implies:

If $\vdash \Box(\psi \equiv \chi)$, then $\vdash \theta \equiv \theta'$

It follows from this and (ϑ) that $\vdash \theta \equiv \theta'$. So by the Rule of Actualization (135), we have $\vdash \mathcal{A}(\theta \equiv \theta')$. But since (139.5) is a theorem, we know $\vdash \mathcal{A}(\theta \equiv \theta') \equiv (\mathcal{A}\theta \equiv \mathcal{A}\theta')$. So by biconditional syllogism (89.3.a), we have $\vdash \mathcal{A}\theta \equiv \mathcal{A}\theta'$, i.e., $\vdash \varphi \equiv \varphi'$.

Inductive Case 6. $\varphi = \Box\theta$. Then φ', i.e., $(\Box\theta)'$, must be $\Box\theta'$. Our IH implies:

If $\vdash \Box(\psi \equiv \chi)$, then $\vdash \theta \equiv \theta'$

It follows from this and (ϑ) that $\vdash \theta \equiv \theta'$. Since no $\star$-theorems were cited to establish this theorem, it follows by the Rule of Necessitation that $\vdash \Box(\theta \equiv \theta')$. But by our proof of (158.6), we know:

$$\vdash \Box(\theta \equiv \theta') \rightarrow (\Box\theta \equiv \Box\theta')$$

Hence it follows by (63.6) that $\vdash \Box\theta \equiv \Box\theta'$, i.e., $\varphi \equiv \varphi'$. ⋈

(159.3) By hypothesis, $\vdash_\Box \psi \equiv \chi$. Then, by RN, $\vdash \Box(\psi \equiv \chi)$. So, where φ' is the result of substituting the formula χ for zero or more occurrences of ψ where the latter is a subformula of φ, it follows by by (159.2) that $\vdash \varphi \equiv \varphi'$. ⋈

(159.4) By hypothesis, $\psi \equiv_{df} \chi$. Then by Rule $\equiv$Df (90.1) of Equivalence by Definition, the $\psi \equiv \chi$ is a theorem. So by RN, $\Box(\psi \equiv \chi)$ is a theorem. If φ' is the result of substituting ψ for zero or more occurrences of the χ where the latter is a subformula of φ, then by (159.2), $\vdash \varphi \equiv \varphi'$ is a theorem.[441] ⋈

(160.1) By hypothesis, $\vdash \Box(\psi \equiv \chi)$ and φ' is the result of substituting the formula χ for zero or more occurrences of ψ where the latter is a subformula of φ. Then, by (159.2), $\vdash \varphi \equiv \varphi'$. By the definition of $\equiv$ and &E, it follows both that $\vdash \varphi \rightarrow \varphi'$ and $\vdash \varphi' \rightarrow \varphi$. So by (63.10), we know both:

(a) $\varphi \vdash \varphi'$

(b) $\varphi' \vdash \varphi$

[441] To justify the $\vdash_\Box$ form of the present rule, we use the $\vdash_\Box$ form of (90.1), which lets us conclude that $\vdash_\Box \psi \equiv \chi$ from the definition $\psi \equiv_{df} \chi$. So by the $\vdash_\Box$ version of (159.3), it follows that $\vdash_\Box \varphi \equiv \varphi'$.

Now we show both directions of the rule. ($\rightarrow$) Assume $\Gamma \vdash \varphi$. Then from this and (a) it follows that $\Gamma \vdash \varphi'$, by (63.8). ($\leftarrow$) By analogous reasoning from (b). $\bowtie$

(160.2) By hypothesis, $\vdash_{\Box} \psi \equiv \chi$ and φ' is the result of substituting the formula χ for zero or more occurrences of ψ where the latter is a subformula of φ. Then, by Rule RN, $\vdash \Box(\psi \equiv \chi)$. Hence by (160.1), $\Gamma \vdash \varphi$ if and only if $\Gamma \vdash \varphi'$. $\bowtie$

(160.3) (Exercise)

(162.1) (Exercise)

(162.2) We reason using a Rule of Necessary Equivalence (159.3):

$\Diamond(\varphi \vee \psi)$	$\equiv$	$\Diamond\neg(\neg\varphi \,\&\, \neg\psi)$	by (88.5.b) and (159.3)
	$\equiv$	$\neg\Box(\neg\varphi \,\&\, \neg\psi)$	by (158.11)
	$\equiv$	$\neg(\Box\neg\varphi \,\&\, \Box\neg\psi)$	by (158.3) and (159.3)
	$\equiv$	$\neg(\neg\Diamond\varphi \,\&\, \neg\Diamond\psi$	by (162.1) and (159.3)
	$\equiv$	$\neg\neg(\Diamond\varphi \vee \Diamond\psi)$	by (88.5.b) and (159.3)
	$\equiv$	$\Diamond\varphi \vee \Diamond\psi$	by (88.3.b) $\bowtie$

(162.3) The tautologies of conjunction simplification (88.1) are $(\varphi \,\&\psi) \rightarrow \varphi$ and $(\varphi \,\&\psi) \rightarrow \psi$. Since these are modally strict theorems, it follows by RM$\Diamond$ (157.2) that $\Diamond(\varphi \,\&\psi) \rightarrow \Diamond\varphi$ and $\Diamond(\varphi \,\&\psi) \rightarrow \Diamond\psi$. So assume $\Diamond(\varphi \,\&\psi)$, for conditional proof. Then both $\Diamond\varphi$ and $\Diamond\psi$. So by &I, we have $\Diamond\varphi \,\&\, \Diamond\psi$. $\bowtie$

(162.4) We reason using a Rule of Necessary Equivalence (159.3):

$\Diamond(\varphi \rightarrow \psi)$	$\equiv$	$\Diamond(\neg\varphi \vee \psi)$	by (88.1.c) and (159.3)
	$\equiv$	$\Diamond\neg\varphi \vee \Diamond\psi$	by (162.2)
	$\equiv$	$\neg\Box\varphi \vee \Diamond\psi$	by (158.1) and (159.3)
	$\equiv$	$\Box\varphi \rightarrow \Diamond\psi$	by (88.1.c) $\bowtie$

(162.5) We reason using a Rule of Necessary Equivalence (159.4):

$\Diamond\Diamond\varphi$	$\equiv$	$\Diamond\neg\Box\neg\varphi$	by definition $\Diamond$ (18.5) and (159.4)
	$\equiv$	$\neg\Box\Box\neg\varphi$	by (158.11) $\bowtie$

(162.6) Assume $\Box(\varphi \vee \psi)$. Suppose, for reductio, $\neg(\Box\varphi \vee \Diamond\psi)$. Then by a De Morgan's law (88.5.d) and &E, it follows both that $\neg\Box\varphi$ and $\neg\Diamond\psi$. By (158.11), the first implies:

(ϑ) $\Diamond\neg\varphi$

But our initial assumption implies, by definition of $\vee$ and the Rule of Substitution for Defined Formulas (160.3), that $\Box(\neg\varphi \rightarrow \psi)$. This latter implies, by the K$\Diamond$ schema (158.13), $\Diamond\neg\varphi \to \Diamond\psi$. From this and ($\vartheta$), it follows that $\Diamond\psi$. Contradiction. ⋈

(162.7) Assume $\Box(\varphi \vee \psi)$ and $\Diamond\neg\varphi$. From the first assumption and (162.6), it follows that $\Box\varphi \vee \Diamond\psi$. From the second assumption it follows that $\neg\Box\varphi$, by (158.11). From our last two results, it follows that $\Diamond\psi$, by disjunctive syllogism (86.4.b). ⋈

(163.1) As an instance of the T schema (45.2), we know $\Box\neg\varphi \to \neg\varphi$. So by contraposition (80), $\neg\neg\varphi \to \neg\Box\neg\varphi$. But $\varphi \to \neg\neg\varphi$ is a tautology (77.2). So by hypothetical syllogism (76.3), it follows that $\varphi \to \neg\Box\neg\varphi$.By the definition of $\Diamond$ (18.5) and the Rule of Definition by Equivalence (72), we independently know $\neg\Box\neg\varphi \to \Diamond\varphi$. So our last two results imply $\varphi \to \Diamond\varphi$, by hypothetical syllogism. ⋈

(163.2) Assume $\Diamond\Box\varphi$. From this and (158.14), it follows by biconditional syllogism that $\neg\Box\Diamond\neg\varphi$. But note that the following is an instance of the 5 schema: $\Diamond\neg\varphi \to \Box\Diamond\neg\varphi$. So by a rule of modus tollens (79.1), $\neg\Diamond\neg\varphi$, which by (158.12), yields $\Box\varphi$. ⋈

(164.1) As an instance of theorem (88.3.b), we know $\mathcal{A}\varphi \equiv \neg\neg\mathcal{A}\varphi$. Independently, a necessary axiom of actuality (44.1) yields, by the commutativity of $\equiv$, the modally strict theorem $\neg\mathcal{A}\varphi \equiv \mathcal{A}\neg\varphi$. Hence, it follows from these two results, by a Rule of Substitution (160.2), that $\mathcal{A}\varphi \equiv \neg\mathcal{A}\neg\varphi$. ⋈

(164.2) As an instance of axiom (46.2), we know $\Box\neg\varphi \equiv \mathcal{A}\Box\neg\varphi$. By a classical tautology (88.4.b), it follows that $\neg\Box\neg\varphi \equiv \neg\mathcal{A}\Box\neg\varphi$. Since $\Diamond\varphi \equiv \neg\Box\neg\varphi$ follows from definition (18.5) by Rule $\equiv$Df (90.1), it follows from the last two results that $\Diamond\varphi \equiv \neg\mathcal{A}\Box\neg\varphi$, by biconditional syllogism. But as an instance of necessary axiom (44.1), we know $\neg\mathcal{A}\Box\neg\varphi \equiv \mathcal{A}\neg\Box\neg\psi$. So again by biconditional syllogism, $\Diamond\varphi \equiv \mathcal{A}\neg\Box\neg\varphi$. But since $\Diamond\varphi \equiv \neg\Box\neg\varphi$ is a modally strict equivalence arising from a definition, it follows by a Rule of Substitution (160.3) that $\Diamond\varphi \equiv \mathcal{A}\Diamond\varphi$. ⋈

(164.3) As an instance of (132), we know $\Box\neg\varphi \to \mathcal{A}\neg\varphi$. So by contraposition, $\neg\mathcal{A}\neg\varphi \to \neg\Box\neg\varphi$. But by (164.1), we know $\mathcal{A}\varphi \to \neg\mathcal{A}\neg\varphi$. So by biconditional syllogism, $\mathcal{A}\varphi \to \neg\Box\neg\varphi$. But since $\neg\Box\neg\varphi \to \Diamond\varphi$, by definition of $\Diamond$ and the Rule of Definition by Equivalence (72), it follows that $\mathcal{A}\varphi \to \Diamond\varphi$. ⋈

(164.4) ($\to$) This is an instance of the the T$\Diamond$ schema (163.1). ($\leftarrow$) Assume $\Diamond\mathcal{A}\varphi$. Then by definition of $\Diamond$, we know $\neg\Box\neg\mathcal{A}\varphi$. But by commuting (44.1), we know $\neg\mathcal{A}\varphi \equiv \mathcal{A}\neg\varphi$ is a modally strict theorem. So it follows by a Rule of Substitution (160.2) that:

(ϑ) $\neg\Box\mathcal{A}\neg\varphi$

Now assume, for reductio, $\neg\mathcal{A}\varphi$. Then by the right-to-left direction of axiom (44.1), $\mathcal{A}\neg\varphi$. But, as an instance of (46.1), we know $\mathcal{A}\neg\varphi \to \Box\mathcal{A}\neg\varphi$. Hence, $\Box\mathcal{A}\neg\varphi$, which contradicts ($\vartheta$). ⋈

(164.5) By the right-to-left direction of (164.4), we know $\Diamond\mathcal{A}\varphi \to \mathcal{A}\varphi$. Independently, by (164.3) we know $\mathcal{A}\varphi \to \Diamond\varphi$. And independently, by (164.2), we know $\Diamond\varphi \to \mathcal{A}\Diamond\varphi$. Hence by hypothetical syllogism, $\Diamond\mathcal{A}\varphi \to \mathcal{A}\Diamond\varphi$. ⋈

(165.1) ($\to$) $\Diamond\varphi \to \Box\Diamond\varphi$ is the 5 schema. ($\leftarrow$) $\Box\Diamond\varphi \to \Diamond\varphi$ is an instance of the T schema (45.2). ⋈

(165.2) ($\to$) $\Box\varphi \to \Diamond\Box\varphi$ is an instance of the T$\Diamond$ schema (163.1). ($\leftarrow$) $\Diamond\Box\varphi \to \Box\varphi$ is an instance of the 5$\Diamond$ schema (163.2) ⋈

(165.3) By the T$\Diamond$ schema (163.1), we know $\varphi \to \Diamond\varphi$. And as an instance of the 5 schema (45.3), we know: $\Diamond\varphi \to \Box\Diamond\varphi$. So by hypothetical syllogism (76.3), it follows that $\varphi \to \Box\Diamond\varphi$. ⋈

(165.4) As an instance of the B schema (165.3), we have $\neg\varphi \to \Box\Diamond\neg\varphi$. It follows from this by contraposition that $\neg\Box\Diamond\neg\varphi \to \neg\neg\varphi$. Since $\neg\neg\varphi \to \varphi$ (77.1), it follows by hypothetical syllogism that:

(ϑ) $\neg\Box\Diamond\neg\varphi \to \varphi$

Now, independently, it follows *a fortiori* from (158.14) that $\Diamond\Box\varphi \to \neg\Box\Diamond\neg\varphi$. From this and ($\vartheta$), it follows by hypothetical syllogism that $\Diamond\Box\varphi \to \varphi$. ⋈

(165.5) $\Box\varphi \to \Box\Diamond\Box\varphi$ is an instance of the B schema (165.3). Independently, since $\Diamond\Box\varphi \to \Box\varphi$ is a $\Box$-theorem (163.2), it follows by RM (157.1) that $\Box\Diamond\Box\varphi \to \Box\Box\varphi$. So by hypothetical syllogism, $\Box\varphi \to \Box\Box\varphi$. ⋈

(165.6) (Exercise)

(165.7) As an instance of (165.5) we have $\Box\neg\varphi \to \Box\Box\neg\varphi$. By a rule of contraposition, this implies $\neg\Box\Box\neg\varphi \to \neg\Box\neg\varphi$. But from the definition of $\Diamond$ (18.5), we may infer that $\neg\Box\neg\varphi \to \Diamond\varphi$ (72). So $\neg\Box\Box\neg\varphi \to \Diamond\varphi$, by hypothetical syllogism. But the left-to-right direction of (162.5) is $\Diamond\Diamond\varphi \to \neg\Box\Box\neg\varphi$. From this and the previous result, it follows that $\Diamond\Diamond\varphi \to \Diamond\varphi$, by hypothetical syllogism. ⋈

(165.8) (Exercise)

(165.9) ($\to$) As an instance of (162.6), we know:

$$\Box(\varphi \lor \Box\psi) \to (\Box\varphi \lor \Diamond\Box\psi)$$

Since the commuted form of (165.2) establishes a modally strict equivalence between $\Diamond\Box\psi$ and $\Box\psi$, the Rule of Substitution (160.2) allows us to infer the following: $\Box(\varphi \lor \Box\psi) \to (\Box\varphi \lor \Box\psi)$. ($\leftarrow$) As an instance of (158.15), we know:

$$(\Box\varphi \lor \Box\Box\psi) \to \Box(\varphi \lor \Box\psi)$$

Since (165.6) establishes a modally strict equivalence between $\Box\Box\psi$ and $\Box\psi$, the Rule of Substitution (160.2) allows us to infer the following: $(\Box\varphi \lor \Box\psi) \to \Box(\varphi \lor \Box\psi)$. ⋈

(165.10) (Exercise)

(165.11) We reason using a Rule of Necessary Equivalence (159.3) as follows:

$$\begin{array}{rcll}
\Diamond(\varphi \,\&\, \Diamond\psi) & \equiv & \Diamond\neg(\neg\varphi \vee \neg\Diamond\psi) & \text{by (88.5.a) and (159.3)} \\
& \equiv & \Diamond\neg(\neg\varphi \vee \Box\neg\psi) & \text{by (162.1) and (159.3)} \\
& \equiv & \neg\Box(\neg\varphi \vee \Box\neg\psi) & \text{by (158.11)} \\
& \equiv & \neg(\Box\neg\varphi \vee \Box\neg\psi) & \text{by (165.9) and (88.4.c)} \\
& \equiv & \neg(\neg\Diamond\varphi \vee \neg\Diamond\psi) & \text{by (162.1) and (159.3) } (\times 2) \\
& \equiv & \Diamond\varphi \,\&\, \Diamond\psi & \text{by (88.5.a)} \qquad \bowtie
\end{array}$$

(165.12) (Exercise)

(165.13) ($\rightarrow$) For the left-to-right direction, our proof strategy is as follows:

(a) Show $\Box(\varphi \rightarrow \Box\psi) \rightarrow (\Diamond\varphi \rightarrow \psi)$, by a modally strict proof.

(b) Conclude $\Box\Box(\varphi \rightarrow \Box\psi) \rightarrow \Box(\Diamond\varphi \rightarrow \psi)$ from (a) by Rule RM (157.1)

(c) Show that the left-to-right direction of our theorem, i.e., $\Box(\varphi \rightarrow \Box\psi) \rightarrow \Box(\Diamond\varphi \rightarrow \psi)$, follows from (b).

It remains to show (a) and (c).

For (a), assume $\Box(\varphi \rightarrow \Box\psi)$. Then by K$\Diamond$ (158.13), it follows that $\Diamond\varphi \rightarrow \Diamond\Box\psi$. But the B$\Diamond$ schema (165.4) is $\Diamond\Box\psi \rightarrow \psi$. So $\Diamond\varphi \rightarrow \psi$ follows by hypothetical syllogism from our last two results.

For (c), assume $\Box(\varphi \rightarrow \Box\psi)$. Then by the 4 schema (165.5), it follows that $\Box\Box(\varphi \rightarrow \Box\psi)$. So by (b), it follows that $\Box(\Diamond\varphi \rightarrow \psi)$.

($\leftarrow$) We leave the right-to-left direction as an exercise. $\bowtie$

(166.1) [We prove only the stronger version.] Assume $\Gamma \vdash_\Box \Diamond\varphi \rightarrow \psi$, i.e., that there is a modally-strict derivation of $\Diamond\varphi \rightarrow \psi$ from Γ. So by the strong form of Rule RM (157.1), it follows that $\Box\Gamma \vdash_\Box \Box\Diamond\varphi \rightarrow \Box\psi$. By (165.3), the instances of the B schema are modally strict theorems, so by (63.3) we have $\Box\Gamma \vdash_\Box \varphi \rightarrow \Box\Diamond\varphi$. Hence, by the $\vdash_\Box$ version of (76.1), it follows that $\Box\Gamma \vdash_\Box \varphi \rightarrow \Box\psi$. $\bowtie$

(166.2) [We prove only the stronger version.] Assume $\Gamma \vdash_\Box \varphi \rightarrow \Box\psi$. Then by the strong form of RM$\Diamond$ (157.2), it follows that $\Box\Gamma \vdash_\Box \Diamond\varphi \rightarrow \Diamond\Box\psi$. But the schema B$\Diamond$ (165.4) is a modally strict theorem, and so $\Box\Gamma \vdash_\Box \Diamond\Box\psi \rightarrow \psi$, by (63.3). Hence, by the $\vdash_\Box$ version of (76.1) it follows that $\Box\Gamma \vdash_\Box \Diamond\varphi \rightarrow \psi$. $\bowtie$

(167.1) We reason as follows:

1.	$\forall\alpha\Box\varphi \rightarrow \Box\varphi$	instance of (95.3)
2.	$\Diamond\forall\alpha\Box\varphi \rightarrow \Diamond\Box\varphi$	from (1) by RM$\Diamond$ (157.2)
3.	$\Diamond\Box\varphi \rightarrow \varphi$	instance of B$\Diamond$ (165.4)
4.	$\Diamond\forall\alpha\Box\varphi \rightarrow \varphi$	from (2),(3) by (76.3)
5.	$\forall\alpha(\Diamond\forall\alpha\Box\varphi \rightarrow \varphi)$	from (5) by GEN
6.	$\forall\alpha(\Diamond\forall\alpha\Box\varphi \rightarrow \varphi) \rightarrow (\Diamond\forall\alpha\Box\varphi \rightarrow \forall\alpha\varphi)$	instance of (95.2)
7.	$\Diamond\forall\alpha\Box\varphi \rightarrow \forall\alpha\varphi$	from (5),(6) by MP
8.	$\forall\alpha\Box\varphi \rightarrow \Box\forall\alpha\varphi$	from (7), by Rule (166.1) $\bowtie$

(167.2) Theorem (95.3) asserts $\forall\alpha\varphi \rightarrow \varphi$. So by Rule RM (157.1), it follows that $\Box\forall\alpha\varphi \rightarrow \Box\varphi$. By GEN, it follows that $\forall\alpha(\Box\forall\alpha\varphi \rightarrow \Box\varphi)$. But since α isn't free in $\Box\forall\alpha\varphi$, it follows by an appropriate instance of (95.2) that $\Box\forall\alpha\varphi \rightarrow \forall\alpha\Box\varphi$. $\bowtie$

(167.3) Given (167.1) and (167.2), it is a modally strict theorem that $\forall\alpha\Box\varphi \equiv \Box\forall\alpha\varphi$. So as an instance, we know $\forall\alpha\Box\neg\varphi \equiv \Box\forall\alpha\neg\varphi$, which commutes to:

(ϑ) $\Box\forall\alpha\neg\varphi \equiv \forall\alpha\Box\neg\varphi$

Hence we may reason using the Rules of Substitution (160.2) and (160.3) as follows:

$\Diamond\exists\alpha\varphi$	$\rightarrow$	$\neg\Box\neg\exists\alpha\varphi$	by definition $\Diamond$ and (72)
	$\rightarrow$	$\neg\Box\forall\alpha\neg\varphi$	by (103.4) and (160.2)
	$\rightarrow$	$\neg\forall\alpha\Box\neg\varphi$	by (ϑ) and (160.2)
	$\rightarrow$	$\exists\alpha\neg\Box\neg\varphi$	by (103.2)
	$\rightarrow$	$\exists\alpha\Diamond\varphi$	by definition $\Diamond$ and (160.3) $\bowtie$

(167.4) Given (167.1) and (167.2), it is a modally strict theorem that $\forall\alpha\Box\varphi \equiv \Box\forall\alpha\varphi$. So as an instance, we know:

(ϑ) $\forall\alpha\Box\neg\varphi \equiv \Box\forall\alpha\neg\varphi$

Hence we may reason using the Rules of Substitution (160.2) and (160.3) as follows:

$\exists\alpha\Diamond\varphi$	$\rightarrow$	$\neg\forall\alpha\neg\Diamond\varphi$	by definition $\exists$ and (72)
	$\rightarrow$	$\neg\forall\alpha\Box\neg\varphi$	by (162.1) and (160.2)
	$\rightarrow$	$\neg\Box\forall\alpha\neg\varphi$	by (ϑ) and (160.2)
	$\rightarrow$	$\Diamond\neg\forall\alpha\neg\varphi$	by (158.11)
	$\rightarrow$	$\Diamond\exists\alpha\varphi$	by definition $\exists$ and (160.3) $\bowtie$

(168.1) Assume $\exists\alpha\Box\varphi$. Now assume τ be an arbitrary such α, so that we know:

(ϑ) $\Box\varphi^{\tau}_{\alpha}$

Since τ is an arbitrarily chosen, primitive constant that is substitutable for, and has the same type as, the variable α in φ, we know independently that $\tau\downarrow$ by (39.2). Hence, by (106):

(ξ) $\Box\tau\downarrow$

Also, independently, Rule $\exists$I (101.1) tells us $\varphi^{\tau}_{\alpha}, \tau\downarrow \vdash \exists\alpha\varphi$. Hence, by RN, $\Box\varphi^{\tau}_{\alpha}, \Box\tau\downarrow \vdash \Box\exists\alpha\varphi$. So from ($\vartheta$) and ($\xi$) it follows that $\Box\exists\alpha\varphi$. By Rule $\exists$E (102), we may discharge (ϑ) and we've thereby derived $\Box\exists\alpha\varphi$ from our initial assumption $\exists\alpha\Box\varphi$. ⋈

(168.2) Theorem (95.3) asserts $\forall\alpha\varphi \to \varphi$. By RM$\Diamond$ (157.2), then, it follows that $\Diamond\forall\alpha\varphi \to \Diamond\varphi$. So by GEN, it follows that $\forall\alpha(\Diamond\forall\alpha\varphi \to \Diamond\varphi)$. But since α isn't free in $\Diamond\forall\alpha\varphi$, it follows by an appropriate instance of (95.2) that $\Diamond\forall\alpha\varphi \to \forall\alpha\Diamond\varphi$. ⋈

(168.3) From (103.5), by RM$\Diamond$ (157.2). ⋈

(168.4) (Exercise)

(168.5) Assume, for conditional proof:

$$\Box\forall\alpha(\varphi \to \psi) \,\&\, \Box\forall\alpha(\psi \to \chi)$$

Then since a conjunction of necessities implies a necessary conjunction (158.3), it follows that:

(ϑ) $\Box(\forall\alpha(\varphi \to \psi) \,\&\, \forall\alpha(\psi \to \chi))$

Note, independently, that the following is an instance of (99.9):

$$(\forall\alpha(\varphi \to \psi) \,\&\, \forall\alpha(\psi \to \chi)) \to \forall\alpha(\varphi \to \chi)$$

Since this is a modally strict theorem, it follows by Rule RM (157.1) that:

$$\Box(\forall\alpha(\varphi \to \psi) \,\&\, \forall\alpha(\psi \to \chi)) \to \Box\forall\alpha(\varphi \to \chi)$$

And from this last result and (ϑ), it follows by MP that $\Box\forall\alpha(\varphi \to \chi)$. ⋈

(168.6) By reasoning analogous to (168.5) but starting with (99.10) and using (158.6) instead of the K axiom. ⋈

(169.1) Since it is a modally strict theorem (106) that $\tau\downarrow \to \Box\tau\downarrow$, it follows by rule (166.2) that $\Diamond\tau\downarrow \to \tau\downarrow$. ⋈

(169.2) (Exercise)

(169.3) It follows from (169.1) that $\neg\tau\downarrow \to \neg\Diamond\tau\downarrow$, by contraposition. But by the right-to-left direction of (162.1), $\neg\Diamond\tau\downarrow \to \Box\neg\tau\downarrow$. So by hypothetical syllogism, $\neg\tau\downarrow \to \Box\neg\tau\downarrow$. ⋈

(169.4) (Exercise)

(170.1) $(\rightarrow)$ By theorem (125.1), we know $\alpha = \beta \rightarrow \Box\alpha = \beta$. Since this is a $\Box$-theorem, it follows by (166.2) that $\Diamond\alpha = \beta \rightarrow \alpha = \beta$. ⋈

(170.2) By (170.1), $\Diamond\alpha = \beta \rightarrow \alpha = \beta$. By contraposition, $\neg\alpha = \beta \rightarrow \neg\Diamond\alpha = \beta$. But by (162.1), $\neg\Diamond\alpha = \beta \rightarrow \Box\neg\alpha = \beta$. Hence, $\neg\alpha = \beta \rightarrow \Box\neg\alpha = \beta$. Using infix notation, $\alpha \neq \beta \rightarrow \Box\alpha \neq \beta$. ⋈

(170.3) $(\rightarrow)$ By (170.2), $\alpha \neq \beta \rightarrow \Box\alpha \neq \beta$, Since this is a modally strict theorem, we may apply (166.2) to conclude $\Diamond\alpha \neq \beta \rightarrow \alpha \neq \beta$. ⋈

(170.4) – **(170.5)** (Exercises)

(171.1) Assume $\Box(\varphi \rightarrow \Box\varphi)$. Assume, for reductio, $\neg\Box(\neg\varphi \rightarrow \Box\neg\varphi)$, i.e., by (158.11), that $\Diamond\neg(\neg\varphi \rightarrow \Box\neg\varphi)$. Then by the relevant instance of the modally strict theorem (88.1.b) and a Rule of Substitution, it follows that $\Diamond(\neg\varphi \,\&\, \neg\Box\neg\varphi)$. By definition (18.5) and a Rule of Substitution, this implies $\Diamond(\neg\varphi \,\&\, \Diamond\varphi)$. Then by (162.3), it follows that both $\Diamond\neg\varphi$ and $\Diamond\Diamond\varphi$. The former is equivalent to $\neg\Box\varphi$ (158.11); the latter reduces to $\Diamond\varphi$, by 4$\Diamond$ (165.7). From $\Diamond\varphi$ and our initial assumption, it follows that $\Diamond\Box\varphi$ (158.13). So by (165.2), $\Box\varphi$. Contradiction. ⋈

(171.1) [Alternative proof:] Our theorem is a consequence of the following derivability claims, by (63.8):

$$\Box(\varphi \rightarrow \Box\varphi) \vdash_{\Box} \Box\Box(\varphi \rightarrow \Box\varphi)$$

$$\Box\Box(\varphi \rightarrow \Box\varphi) \vdash_{\Box} \Box(\neg\varphi \rightarrow \Box\neg\varphi)$$

The first claim is immediate by the 4 schema (165.5). The second claim follows by RN from $\Box(\varphi \rightarrow \Box\varphi) \vdash_{\Box} (\neg\varphi \rightarrow \Box\neg\varphi)$. So it suffices to show the latter. Assume $\Box(\varphi \rightarrow \Box\varphi)$. Now assume $\neg\varphi$. Then $\Diamond\neg\varphi$, by the T$\Diamond$ schema. So $\neg\Box\varphi$. But from our first assumption, it follows by (172.3) that $\neg\Box\varphi \equiv \Box\neg\varphi$. Hence $\Box\neg\varphi$. ⋈

(171.2) Assume $\Box(\varphi \rightarrow \Box\varphi)$ and $\Box(\psi \rightarrow \Box\psi)$. For reductio, assume:

$$\neg\Box((\varphi \rightarrow \psi) \rightarrow \Box(\varphi \rightarrow \psi))$$

By reasoning analogous to steps in the previous proof, we therefore know both:

(A) $\Diamond(\varphi \rightarrow \psi)$

(B) $\Diamond\neg\Box(\varphi \rightarrow \psi)$

To complete our reductio, we show that (B) implies both $\Diamond\varphi$ and $\neg\Box\psi$, and that from $\Diamond\varphi$ and our initial assumptions, (A) implies $\Box\psi$.

To see that (B) implies $\Diamond\varphi$ and $\neg\Box\psi$, note that by (158.11) and a Rule of Substitution, (B) is equivalent to $\Diamond\Diamond\neg(\varphi \rightarrow \psi)$. This reduces to $\Diamond\neg(\varphi \rightarrow \psi)$, by the 4$\Diamond$ schema (165.7). But this is equivalent to $\Diamond(\varphi \,\&\, \neg\psi)$ (exercise), which implies $\Diamond\varphi$ and $\Diamond\neg\psi$, by (162.3). The latter yields $\neg\Box\psi$.

To see that from $\Diamond\varphi$ and our initial assumptions, (A) implies $\Box\psi$, note that from $\Diamond\varphi$ and our first assumption, $\Box(\varphi \to \Box\varphi)$, it follows that $\Diamond\Box\varphi$, which reduces to $\Box\varphi$. But by (162.4), $\Box\varphi$ and (A) imply $\Diamond\psi$. From $\Diamond\psi$ and our second assumption, $\Box(\psi \to \Box\psi)$, it follows that $\Diamond\Box\psi$, which reduces to $\Box\psi$. Contradiction. $\bowtie$

(171.3) – **(171.5)** (Exercises)

(171.6) Assume $\Box\forall\alpha(\varphi \to \Box\varphi)$ and for reductio, $\neg\Box(\forall\alpha\varphi \to \Box\forall\alpha\varphi)$. Then, by now familiar reasoning, we know both:

(A) $\Diamond\forall\alpha\varphi$

(B) $\Diamond\neg\Box\forall\alpha\varphi$

We complete our reductio by showing that (A) and our initial assumption imply $\Box\forall\alpha\varphi$ whereas (B) implies $\neg\Box\forall\alpha\varphi$.

To see that (A) and our initial assumptions imply $\Box\forall\alpha\varphi$, note first that the following modal closure of an instance of axiom (39.3) is an axiom:

$$\Box(\forall\alpha(\varphi \to \Box\varphi) \to (\forall\alpha\varphi \to \forall\alpha\Box\varphi))$$

Hence by the K axiom, it follows that:

$$\Box\forall\alpha(\varphi \to \Box\varphi) \to \Box(\forall\alpha\varphi \to \forall\alpha\Box\varphi)$$

From this and our initial assumption, it follows that $\Box(\forall\alpha\varphi \to \forall\alpha\Box\varphi)$. From this and (A) it follows by (158.13) that $\Diamond\forall\alpha\Box\varphi$. But by the modally strict equivalence of BF and CBF, by (167.1) and (167.2), it follows by a Rule of Substitution that $\Diamond\Box\forall\alpha\varphi$. Then by (165.2), $\Box\forall\alpha\varphi$.

To see that (B) implies $\neg\Box\forall\alpha\varphi$, note that by the modally strict theorem (158.11) and a Rule of Substitution, (B) implies $\Diamond\Diamond\neg\forall\alpha\varphi$. This reduces to $\Diamond\neg\forall\alpha\varphi$, by the 4$\Diamond$ schema. Hence, $\neg\Box\forall\alpha\varphi$. Contradiction. $\bowtie$

(172.1) ($\to$) Assume $\Box(\varphi \to \Box\varphi)$. Then by (158.13), $\Diamond\varphi \to \Diamond\Box\varphi$. But by the 5$\Diamond$ schema (163.2), $\Diamond\Box\varphi \to \Box\varphi$. Hence, by hypothetical syllogism, $\Diamond\varphi \to \Box\varphi$.[442]
($\leftarrow$) Assume $\Diamond\varphi \to \Box\varphi$. For reductio, assume $\neg\Box(\varphi \to \Box\varphi)$, i.e., by (158.11), assume:

(ϑ) $\Diamond\neg(\varphi \to \Box\varphi)$

But it is a modally strict theorem that $\neg(\varphi \to \Box\varphi) \equiv (\varphi \,\&\, \neg\Box\varphi)$, by (88.1.b). Hence, it follows from (ϑ) and this last fact by the Rule of Substitution (160.2) that $\Diamond(\varphi \,\&\, \neg\Box\varphi)$. So by (162.3) and &E, it follows that both $\Diamond\varphi$ and $\Diamond\neg\Box\varphi$.

[442] Here is an alternative proof of this direction. Assume $\Box(\varphi \to \Box\varphi)$. Then by (165.13), $\Box(\Diamond\varphi \to \varphi)$. Hence by the K axiom, $\Box\Diamond\varphi \to \Box\varphi$. Now assume $\Diamond\varphi$. Then by the 5 axiom (45.3), $\Box\Diamond\varphi$. Hence, $\Box\varphi$.

The former implies, by our initial assumption, $\Box\varphi$. The latter implies, by the modally strict theorem $\neg\Box\varphi \equiv \Diamond\neg\varphi$ and the Rule of Substitution, $\Diamond\Diamond\neg\varphi$, which by the 4◊ principle (165.7) implies $\Diamond\neg\varphi$, i.e, $\neg\Box\varphi$. Contradiction. ⋈

(172.2) Since $\Box(\varphi \to \Box\varphi)$ implies $\Diamond\varphi \to \Box\varphi$ by (172.1), it suffices to show that the latter implies $\Diamond\varphi \equiv \Box\varphi$. But this is easy, since we only have to show $\Box\varphi \to \Diamond\varphi$, which follows from the T and T◊ schemata. ⋈

(172.3) Assume $\Box(\varphi \to \Box\varphi)$. ($\to$) Assume $\neg\Box\varphi$. Our global assumption implies $\Diamond\varphi \to \Box\varphi$, by (172.1). Hence, $\neg\Diamond\varphi$, i.e., $\Box\neg\varphi$, by (162.1). ($\leftarrow$) Assume $\Box\neg\varphi$. Then $\neg\varphi$ by the T schema, and so $\Diamond\neg\varphi$, by the T◊ schema. Hence, $\neg\Box\varphi$, by (158.11). ⋈

(172.4) Assume both the antecedent and the antecedent of the consequent:

(ϑ) $\Box(\varphi \to \Box\varphi) \,\&\, \Box(\psi \to \Box\psi)$

(ξ) $\Box\varphi \equiv \Box\psi$

From (ξ) it follows that $(\Box\varphi \,\&\, \Box\psi) \vee (\neg\Box\varphi \,\&\, \neg\Box\psi)$, by an appropriate instance of (88.4.g). We now reason by cases from the two disjuncts to show, in each case, that $\Box(\varphi \equiv \psi)$:

- Assume $\Box\varphi \,\&\, \Box\psi$. Then it follows from (158.7) that $\Box(\varphi \equiv \psi)$, by a biconditional syllogism (89.3.a).
- Assume $\neg\Box\varphi \,\&\, \neg\Box\psi$. Note independently that by (172.3), the conjuncts of (ϑ) imply, respectively:

  $\neg\Box\varphi \equiv \Box\neg\varphi$

  $\neg\Box\psi \equiv \Box\neg\psi$

  So we may easily derive $\Box\neg\varphi \,\&\, \Box\neg\psi$ from our local assumption. But then by (158.3), it follows that $\Box(\neg\varphi \,\&\, \neg\psi)$, and by (158.9) that $\Box(\varphi \equiv \psi)$. ⋈

(172.5) Though we could follow the proof strategy in (171.2), the following is an alternative strategy:

(A) Show $\Box(\varphi \to \Box\varphi), \Box(\psi \to \Box\psi) \vdash_{\Box} (\varphi \equiv \psi) \to \Box(\varphi \equiv \psi)$

(B) Infer from (A), by Rule RN, that:

$$\Box\Box(\varphi \to \Box\varphi), \Box\Box(\psi \to \Box\psi) \vdash_{\Box} \Box((\varphi \equiv \psi) \to \Box(\varphi \equiv \varphi))$$

(C) Note that by the 4 schema and (63.10), we independently know:

$$\Box(\varphi \to \Box\varphi) \vdash_{\Box} \Box\Box(\varphi \to \Box\varphi)$$
$$\Box(\psi \to \Box\psi) \vdash_{\Box} \Box\Box(\psi \to \Box\psi)$$

(D) Conclude from (C) and (B) that:

$$\Box(\varphi \rightarrow \Box\varphi), \Box(\psi \rightarrow \Box\psi) \vdash_{\Box} \Box((\varphi \equiv \psi) \rightarrow \Box(\varphi \equiv \varphi))$$

(E) Reason from (D), using two applications of the Deduction Theorem (75) and then an application of Importation (88.7.b), to obtain the following modally strict theorem:

$$(\Box(\varphi \rightarrow \Box\varphi) \,\&\, \Box(\psi \rightarrow \Box\psi)) \rightarrow \Box((\varphi \equiv \psi) \rightarrow \Box(\varphi \equiv \psi))$$

Since (B) – (E) are straightforward, it remains to show (A). Our assumptions are:

$\Box(\varphi \rightarrow \Box\varphi)$

$\Box(\psi \rightarrow \Box\psi)$

Then by the T schema, it follows from each, respectively, that:

(ϑ) $\varphi \rightarrow \Box\varphi$

(ξ) $\psi \rightarrow \Box\psi$

Now to complete the proof of (A), we have to show: $(\varphi \equiv \psi) \rightarrow \Box(\varphi \equiv \psi)$. So assume $\varphi \equiv \psi$. Then either $\varphi \,\&\, \psi$ or $\neg\varphi \,\&\, \neg\psi$ (88.4.g) and so we show $\Box(\varphi \equiv \psi)$ in both cases:

$\varphi \,\&\, \psi$. The first conjunct and (ϑ) yield $\Box\varphi$, and the second conjunct and (ξ) yield $\Box\psi$. So $\Box\varphi \,\&\, \Box\psi$, which implies $\Box\varphi \equiv \Box\psi$, by (88.4.g). From this and our assumptions $\Box(\varphi \rightarrow \Box\varphi)$ and $\Box(\psi \rightarrow \Box\psi)$, it follows by (172.4) that $\Box(\varphi \equiv \psi)$.

$\neg\varphi \,\&\, \neg\psi$. Then both $\Diamond\neg\varphi$ and $\Diamond\neg\psi$, i.e., both $\neg\Box\varphi$ and $\neg\Box\psi$. It follows from the first and (ϑ) that $\neg\varphi$ and from the second and (ξ) that $\neg\psi$. Hence $\Diamond\neg\varphi$ and $\Diamond\neg\psi$, i.e., $\neg\Box\varphi$ and $\neg\Box\psi$. Then $\Box\varphi \equiv \Box\psi$, by (88.4.g). So again it follows from this and our assumptions $\Box(\varphi \rightarrow \Box\varphi)$ and $\Box(\psi \rightarrow \Box\psi)$ by (172.4) that $\Box(\varphi \equiv \psi)$. ⋈

(172.6) Assume $\Box(\varphi \rightarrow \Box\varphi)$ and $\varphi \rightarrow \Box\psi$. Assume, for reductio, that $\neg\Box(\varphi \rightarrow \psi)$, i.e., by (158.11), $\Diamond\neg(\varphi \rightarrow \psi)$, i.e., by (88.1.b) and a Rule of Subtitution, $\Diamond(\varphi \,\&\neg\psi)$. So by (162.3), we know both $\Diamond\varphi$ and $\Diamond\neg\psi$. From the former and our first assumption, it follows by K$\Diamond$ (158.13) that $\Diamond\Box\varphi$, which by (163.2) reduces to $\Box\varphi$ and, hence, φ, by the T schema. But from φ, our second assumption implies $\Box\psi$, which contradicts $\Diamond\neg\psi$, i.e., $\neg\Box\psi$. ⋈

(172.7) (Exercise)

(173) This is proved in the text.

(174.1) By commuting (164.4), we know $\Diamond\mathcal{A}\varphi \equiv \mathcal{A}\varphi$. But theorem (139.6) asserts $\mathcal{A}\varphi \equiv \Box\mathcal{A}\varphi$. Hence $\Diamond\mathcal{A}\varphi \equiv \Box\mathcal{A}\varphi$. ⋈

(174.2) Assume $\Box(\varphi \to \Box\varphi)$. Then by (172.1):

(ϑ) $\Diamond\varphi \to \Box\varphi$

($\to$) Assume $\mathcal{A}\varphi$. Then by (164.3), it follows that $\Diamond\varphi$. Hence, by (ϑ), $\Box\varphi$, and so by the T schema (45.2), φ. ($\leftarrow$) Assume φ. Then $\Diamond\varphi$, by the T$\Diamond$ schema (163.1). Hence $\Box\varphi$, by (ϑ). So, $\mathcal{A}\varphi$, by (132). ⋈

(174.3) ($\to$) We know that both of the following are modally strict theorems:

$$\Box(\varphi \to \Box\varphi) \to (\mathcal{A}\varphi \equiv \varphi) \qquad (174.2)$$

$$(\mathcal{A}\varphi \equiv \varphi) \to (\mathcal{A}\varphi \to \varphi) \qquad \text{(by definition of } \equiv \text{ and \&E)}$$

Hence, by biconditional syllogism, it is a modally strict theorem that:

$$\Box(\varphi \to \Box\varphi) \to (\mathcal{A}\varphi \to \varphi)$$

So by Rule RM:

$$\Box\Box(\varphi \to \Box\varphi) \to \Box(\mathcal{A}\varphi \to \varphi)$$

By the relevant instance of the 4 schema, it follows from this by hypothetical syllogism that $\Box(\varphi \to \Box\varphi) \to \Box(\mathcal{A}\varphi \to \varphi)$. ($\leftarrow$) Assume $\Box(\mathcal{A}\varphi \to \varphi)$. Then by the T schema $\mathcal{A}\varphi \to \varphi$. But we can now establish $\varphi \to \mathcal{A}\varphi$ by modally strict reasoning:

> Assume φ and, for reductio, $\neg\mathcal{A}\varphi$. Then by (44.1), $\mathcal{A}\neg\varphi$. But we've established that $\mathcal{A}\varphi \to \varphi$, so if we consider the instance in which we substitute $\neg\varphi$ for φ, it follows from $\mathcal{A}\neg\varphi$ that $\neg\varphi$. Contradiction.

Now by axiom (46.1), we know $\mathcal{A}\varphi \to \Box\mathcal{A}\varphi$. So $\varphi \to \Box\mathcal{A}\varphi$, by hypothetical syllogism. Moreover, it follows from our initial hypothesis by a relevant instance of (158.6) that $\Box\mathcal{A}\varphi \equiv \Box\varphi$, which implies $\Box\mathcal{A}\varphi \to \Box\varphi$. So our last two conditionals imply, by hypothetical syllogism, that $\varphi \to \Box\varphi$. By conditional proof, we've thus established, as a modally strict theorem, that:

$$\Box(\mathcal{A}\varphi \to \varphi) \to (\varphi \to \Box\varphi)$$

So by RM, $\Box\Box(\mathcal{A}\varphi \to \varphi) \to \Box(\varphi \to \Box\varphi)$. Again, by the relevant instance of the 4 schema, it follows by hypothetical syllogism that $\Box(\mathcal{A}\varphi \to \varphi) \to \Box(\varphi \to \Box\varphi)$. ⋈

(174.4) Assume:

(ϑ) $\Box\forall x(\varphi \to \Box\varphi)$

(ξ) $\exists!x\varphi$

We want to show $\imath x\varphi\downarrow$, but by (152.1), it suffices to show: $\exists! x\mathcal{A}\varphi$. Now from ($\xi$), let a be such an object, so that we know:

(ζ) $\varphi^a_x \,\&\, \forall y(\varphi^y_x \to y = a)$

Since we want to show $\exists! x\mathcal{A}\varphi$ and a is going to be our witness, the definition of the uniqueness quantifier requires us to show: $\mathcal{A}\varphi^a_x \,\&\, \forall y(\mathcal{A}\varphi^y_x \to y = a)$. Note independently that by the Converse Barcan Formula (167.2), (ϑ) implies:

(A) $\forall x\Box(\varphi \to \Box\varphi)$

Instantiating to a, it follows that $\Box(\varphi^a_x \to \Box\varphi^a_x)$, and so by the T schema, $\varphi^a_x \equiv \Box\varphi^a_x$. From this and the first conjunct of (ζ), it follows that $\Box\varphi^a_x$, which immediately implies $\mathcal{A}\varphi^a_x$.

Now by GEN, it remains to show $\mathcal{A}\varphi^y_x \to y = a$. So assume $\mathcal{A}\varphi^y_x$. Now independently, it follows from (A) that $\Box(\varphi^y_x \to \Box\varphi^y_x)$. So by (172.1), $\Diamond\varphi^y_x \to \Box\varphi^y_x$. But from our assumption $\mathcal{A}\varphi^y_x$ it follows that $\Diamond\varphi^y_x$. Hence $\Box\varphi^y_x$, and so φ^y_x. Then by the second conjunct of (ζ), $y = a$. So we've established $\exists! x\mathcal{A}\varphi$, which implies $\imath x\varphi\downarrow$, by (152.1). ⋈

(174.5) (Exercise) [Hint: This follows from (174.2) and the modally strict version of Hintikka's scheme (148).]

(174.6) Assume:

(a) $\Box\forall\alpha(\varphi \to \Box\varphi)$

(b) $\exists!\alpha\varphi$

Now pick β to be a variable that doesn't occur free, and is substitutable for α, in φ. Then from (b), it follows by definition (127.1) of the uniqueness quantifier and our Rule $\equiv_{df}$E of Definiendum Elimination (90.2) that:

(ϑ) $\exists\alpha(\varphi \,\&\, \forall\beta(\varphi^\beta_\alpha \to \beta = \alpha))$

So suppose τ is an arbitrary such entity, i.e., suppose:

(ξ) $\varphi^\tau_\alpha \,\&\, \forall\beta(\varphi^\beta_\alpha \to \beta = \tau)$,

where τ is an arbitrarily chosen constant of the same type as α and β. Our strategy is as follows:

(A) Infer $\Box\varphi^\tau_\alpha$ from (ξ) and (a)

(B) Show $\Box\forall\beta(\varphi^\beta_\alpha \to \beta = \tau)$

(C) Conclude $\Box(\varphi^\tau_\alpha \,\&\, \forall\beta(\varphi^\beta_\alpha \to \beta = \tau))$ from (A) and (B)

(D) Conclude $\exists\alpha\Box(\varphi \,\&\, \forall\beta(\varphi^\beta_\alpha \to \beta = \alpha))$ from (C)

(E) Conclude $\Box\exists\alpha(\varphi \;\&\; \forall\beta(\varphi^{\beta}_{\alpha} \to \beta = \alpha))$ from (D) by the Buridan formula (168.1)

(F) Conclude $\Box\exists!\alpha\varphi$ from (E) by the definition of the uniqueness quantifier and a Rule of Substitution.

Since all of the steps are straightforward except for (B), we conclude with a proof of (B). By the Barcan Formula, it suffices to show $\forall\beta\Box(\varphi^{\beta}_{\alpha} \to \beta = \tau)$. Since β isn't free in any assumption, it suffices by GEN to show $\Box(\varphi^{\beta}_{\alpha} \to \beta = \tau)$. For reductio, assume not, i.e., that $\neg\Box(\varphi^{\beta}_{\alpha} \to \beta = \tau)$, i.e., $\Diamond\neg(\varphi^{\beta}_{\alpha} \to \beta = \tau)$. So by a Rule of Substitution, $\Diamond(\varphi^{\beta}_{\alpha} \;\&\; \neg\beta = \tau)$. So by distributing the $\Diamond$ (162.3), we know both:

(c) $\Diamond\varphi^{\beta}_{\alpha}$

(d) $\Diamond\neg\beta = \tau$

But note independently that (a) implies $\forall\alpha\Box(\varphi \to \Box\varphi)$. So instantiating to φ^{β}_{α}, we know: $\Box(\varphi^{\beta}_{\alpha} \to \Box\varphi^{\beta}_{\alpha})$. But by (172.1), this last result implies $\Diamond\varphi^{\beta}_{\alpha} \to \Box\varphi^{\beta}_{\alpha}$. So by (c), $\Box\varphi^{\beta}_{\alpha}$ and hence φ^{β}_{α}. But then by the second conjunct of (ξ), $\beta = \tau$. So by by (125.2), $\Box\beta = \tau$, which contradicts (d). ⋈

(175.1) Let α, β be variables of the same type. By (125.1), we know $\alpha = \beta \to \Box\alpha = \beta$. Since this is a modally strict theorem, it follows by RN that $\Box(\alpha = \beta \to \Box\alpha = \beta)$. So, by (174.2), it follows that $\mathcal{A}\alpha = \beta \equiv \alpha = \beta$. So by commutativity of $\equiv$, $\alpha = \beta \equiv \mathcal{A}\alpha = \beta$. ⋈

(175.2) We may reason as follows:

$$\begin{array}{rcll} \alpha \neq \beta & \equiv & \neg\alpha = \beta & \text{by (24) and Rule} \equiv\text{Df} \\ & \equiv & \neg\mathcal{A}\alpha = \beta & \text{by (175.1) and (88.4.b)} \\ & \equiv & \mathcal{A}\neg\alpha = \beta & \text{commute (44.1)} \\ & \equiv & \mathcal{A}\alpha \neq \beta & \text{(24), Rule} \equiv\text{Df, (159.3)} \quad ⋈ \end{array}$$

(176.1) We want to establish:

$$\mathcal{A}\exists!\alpha\varphi \equiv \exists!\alpha\mathcal{A}\varphi$$

By a fact about the uniqueness quantifier (127.2), we have to show, for some variable β that is substitutable for α in φ and that doesn't occur free in φ:

$$\mathcal{A}\exists\alpha\forall\beta(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha) \equiv \exists\alpha\forall\beta(\mathcal{A}\varphi^{\beta}_{\alpha} \equiv \beta = \alpha)$$

We do this by first noting some modally strict theorems:

(ϑ) $\mathcal{A}\forall\beta(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha) \equiv \forall\beta\mathcal{A}(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha)$ (44.3)

(ξ) $\mathcal{A}(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha) \equiv (\mathcal{A}\varphi^{\beta}_{\alpha} \equiv \mathcal{A}\beta = \alpha)$ (139.5)

(ζ) $\mathcal{A}\beta = \alpha \equiv \beta = \alpha$ (175.1)

Hence we may reason as follows:

$$\begin{aligned}
\mathcal{A}\exists\alpha\forall\beta(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha) &\equiv \exists\alpha\mathcal{A}\forall\beta(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha) && \text{instance of (139.10)}\\
&\equiv \exists\alpha\forall\beta\mathcal{A}(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha) && \text{by } (\vartheta), (159.3)\\
&\equiv \exists\alpha\forall\beta(\mathcal{A}\varphi^{\beta}_{\alpha} \equiv \mathcal{A}\beta = \alpha) && \text{by } (\xi), (159.3)\\
&\equiv \exists\alpha\forall\beta(\mathcal{A}\varphi^{\beta}_{\alpha} \equiv \beta = \alpha) && \text{by } (\zeta), (159.3) \qquad \bowtie
\end{aligned}$$

(176.2) By (152.1) we know $\imath x\varphi\downarrow \equiv \exists!x\mathcal{A}\varphi$. But if we set the variable α in (176.1) to x, we obtain the instance $\mathcal{A}\exists!x\varphi \equiv \exists!x\mathcal{A}\varphi$, which commutes to $\exists!x\mathcal{A}\varphi \equiv \mathcal{A}\exists!x\varphi$. Hence $\imath x\varphi\downarrow \equiv \mathcal{A}\exists!x\varphi$. $\bowtie$

(177.1) By definition of $\downarrow$ (20.1) and Rule $\equiv_{df}$I, we have to show $\exists F F\imath x(x = y)$. Consider the well-defined property $L =_{df} [\lambda x\, E!x \to E!x]$. By $\exists$I, it suffices to show $L\imath x(x = y)$. By the modally strict version of Russell's analysis (151), we have to show:

$$\exists x(\mathcal{A}x = y \,\&\, \forall z(\mathcal{A}z = y \to z = x) \,\&\, Lx)$$

But again by $\exists$I and the fact that $y\downarrow$ (39.2), it suffices to show that y is a witness to this claim, i.e., to show:

(ϑ) $\mathcal{A}y = y \,\&\, \forall z(\mathcal{A}z = y \to z = y) \,\&\, Ly$

Before we begin, we establish some consequences of (175.1), which has as an instance that $x = y \equiv \mathcal{A}x = y$. By a single application of GEN, it follows that $\forall x(x = y \equiv \mathcal{A}x = y)$. Since both $y\downarrow$ and $z\downarrow$, we may infer both of the following from this last result:

(ξ) $y = y \equiv \mathcal{A}y = y$

(ζ) $z = y \equiv \mathcal{A}z = y$

The first conjunct of (ϑ) follows from (ξ) and $y = y$, which we know by Rule =I. To show the second conjunct of (ϑ), it suffices, by GEN, to show $\mathcal{A}z = y \to z = y$. But this is just the right-to-left direction of (ζ). And the third conjunct of (ϑ) follows from the fact that $\forall zLz$, which we know by reasoning from the definition of L (exercise). $\bowtie$

(177.2) Since the closures of axiom (47) are axioms, the following is an axiom:

> $\forall x(x = \imath x\varphi \equiv \forall z(\mathcal{A}\varphi^{z}_{x} \equiv z = x))$, provided z is substitutable for x in φ and doesn't occur free in φ

Now consider the variable y. We not only know $y\downarrow$, but also that no matter what φ is, y is substitutable for x in the matrix of the above universal claim, since the matrix has only two free occurrences of x, neither of which is in the scope of a variable-binding operator that binds y. So we may instantiate that claim to the variable y and obtain:

$y = \imath x\varphi \equiv \forall z(\mathcal{A}\varphi^z_x \equiv z = y)$, provided z is substitutable for x in φ and doesn't occur free in φ

Now let φ be the formula $x = y$. Then z is substitutable for x in φ and doesn't occur free in φ. So as an instance of our last result, we know:

(ϑ) $y = \imath x(x = y) \equiv \forall z(\mathcal{A}z = y \equiv z = y)$

Now, independently, by (175.1), we know $x = y \equiv \mathcal{A}x = y$, which commutes to $\mathcal{A}x = y \equiv x = y$. By GEN, it follows that $\forall x(\mathcal{A}x = y \equiv x = y)$. By the Rule of Alphabetic Variants, it follows that:

$$\forall z(\mathcal{A}z = y \equiv z = y)$$

But this last fact and (ϑ) jointly imply $y = \imath x(x = y)$. ⋈

(178.1) Consider any $n \geq 1$ and assume $x_1 \ldots x_n F^n$. Then by axiom (50), biconditional syllogism, and &E, we know all of the following:

$$x_1[\lambda y\, F^n y x_2 \ldots x_n]$$
$$x_2[\lambda y\, F^n x_1 y x_3 \ldots x_n]$$
$$\vdots$$
$$x_n[\lambda y\, F^n x_1 \ldots x_{n-1} y]$$

By (106), $F{\downarrow} \rightarrow \Box F{\downarrow}$, and by axiom (51), each of these unary encoding claims implies its own necessitation. Hence, we know:

$$\Box x_1[\lambda y\, F^n y x_2 \ldots x_n]$$
$$\Box x_2[\lambda y\, F^n x_1 y x_3 \ldots x_n]$$
$$\vdots$$
$$\Box x_n[\lambda y\, F^n x_1 \ldots x_{n-1} y]$$

Now we leave it as an exercise to prove the generalized version of theorem (158.3), i.e., to prove $\Box(\varphi_1 \,\&\, \ldots \,\&\, \varphi_n) \equiv (\Box\varphi_1 \,\&\, \ldots \,\&\, \Box\varphi_n)$. So, after we conjoin last set of lines displayed above by $n-1$ applications of &I, it follows that:

(ϑ) $\Box(x_1[\lambda y\, F^n y x_2 \ldots x_n] \,\&\, x_2[\lambda y\, F^n x_1 y x_3 \ldots x_n] \,\&\, \ldots \,\&\, x_n[\lambda xy\, F^n x_1 \ldots x_{n-1} y])$

Independently, if we apply Rule RE to the necessary axiom (50), then we obtain:

(ξ) $\Box x_1 \ldots x_n F^n \equiv$
$\Box(x_1[\lambda y\, F^n y x_2 \ldots x_n] \,\&\, x_2[\lambda y\, F^n x_1 y x_3 \ldots x_n] \,\&\, \ldots \,\&\, x_n[\lambda y\, F^n x_1 \ldots x_{n-1} y])$

From (ξ) and (ϑ), it follows by biconditional syllogism (89.3.b) that $\Box x_1 \ldots x_n F^n$. ⋈

(178.2) Assume $\neg x_1 \ldots x_n F^n$, and for reductio, assume $\neg\Box\neg x_1 \ldots x_n F^n$. Then by definition of $\Diamond$ and Rule $\equiv$Df (90.1), it follows from the latter that $\Diamond x_1 \ldots x_n F^n$. But, independently, it follows from the modally strict theorem $x_1 \ldots x_n F^n \to \Box x_1 \ldots x_n F^n$ (178.1) that $\Diamond x_1 \ldots x_n F^n \to x_1 \ldots x_n F^n$, by (166.2). Hence $x_1 \ldots x_n F^n$. Contradiction. ⋈

(179.1) Since theorem (178.1), i.e., $x_1 \ldots x_n F^n \to \Box x_1 \ldots x_n F^n$, is modally strict, it follows by RN that $\Box(x_1 \ldots x_n F^n \to \Box x_1 \ldots x_n F^n)$. By $\vee$I, it follows that:

$$\Box(x_1 \ldots x_n F^n \to \Box x_1 \ldots x_n F^n) \vee (\Diamond x_1 \ldots x_n F^n \to \Box x_1 \ldots x_n F^n)$$

So by theorem (172.2), it follows that $\Diamond x_1 \ldots x_n F^n \equiv \Box x_1 \ldots x_n F^n$. ⋈

(179.2) By (178.1) for ($\to$) direction and the T schema for the ($\leftarrow$) direction. ⋈

(179.3) ($\to$) By the $\Box$-theorem (178.1), i.e., $x_1 \ldots x_n F^n \to \Box x_1 \ldots x_n F^n$, and rule (166.2). ($\leftarrow$) By the T$\Diamond$ (163.1) schema. ⋈

(179.4) To make the proof more easily readable, we prove this for the unary case. For $n \geq 2$, the proof is obtained by analogous reasoning in which appeals to appeals to theorem (178.1) replace appeals to axiom (51).

($\to$) Assume:

(ϑ) $xF \equiv yG$

To show $\Box xF \equiv \Box yG$, we show both directions:

> ($\to$) Assume $\Box xF$. Then by the T schema (45.2), xF. So by (ϑ), yG. Then by axiom (51), $\Box yG$.
>
> ($\leftarrow$) By analogous reasoning.

($\leftarrow$) Assume $\Box xF \equiv \Box yG$. Again, we show both directions:

> ($\to$) Assume xF. Then by axiom (51), $\Box xF$. From this and our global assumption, it follows that $\Box yG$. Hence, by the T schema (45.2), yG.
>
> ($\leftarrow$) By analogous reasoning. ⋈

(179.5) Again, without loss of generality, we prove this for the unary case. For $n \geq 2$, the proof is obtained by analogous reasoning in which appeals to theorem (178.1) replace appeals to axiom (51).

($\to$) This direction is immediate from the relevant instance of theorem (158.6), which asserts that $\Box(\varphi \equiv \psi) \to (\Box\varphi \equiv \Box\psi)$. Alternatively, if one assumes the antecedent, i.e., $\Box(x_1 \ldots x_n F^n \equiv y_1 \ldots y_n G^n)$ and applies the T schema, we obtain $x_1 \ldots x_n F^n \equiv y_1 \ldots y_n G^n$, which by the left-to-right direction of the previous theorem (179.4) implies $\Box x_1 \ldots x_n F^n \equiv \Box y_1 \ldots y_n G^n$.

($\leftarrow$) As an instance of (172.4), we know:

(ϑ) $(\Box(xF \to \Box xF) \,\&\, \Box(yG \to \Box yG)) \to ((\Box xF \equiv \Box yG) \to \Box(xF \equiv yG))$

But $xF \to \Box xF$ and $yG \to \Box yG$ are both just instances of axiom (51). Hence, by RN, we have both $\Box(xF \to \Box xF)$ and $\Box(yG \to \Box yG)$. So it follows from (ϑ) that $(\Box xF \equiv \Box yG) \to \Box(xF \equiv yG)$. ⋈

(179.6) Without loss of generality, we prove only the unary case. By the commutativity of the biconditional (88.2.e), (179.5) converts to $(\Box xF \equiv \Box yG) \equiv \Box(xF \equiv yG)$. Then (179.4) and this last result imply $(xF \equiv yG) \equiv \Box(xF \equiv yG)$, by $\equiv$E (89.3.e). ⋈

(179.7) ($\to$) By theorem (178.2). ($\leftarrow$) By the T schema. ⋈

(179.8) Without loss of generality, we prove only the unary case. Theorem (179.2) is that $xF \equiv \Box xF$. So by a classical tautology (88.4.b), it follows that $\neg xF \equiv \neg\Box xF$. Independently, as an instance of (158.11), we know that $\neg\Box xF \equiv \Diamond\neg xF$. So by the transitivity of $\equiv$ (89.3.e), it follows that $\neg xF \equiv \Diamond\neg xF$, which commutes to $\Diamond\neg xF \equiv \neg xF$. ⋈

(179.9) (Exercise)

(179.10) Since $x_1 \ldots x_n F \to \Box x_1 \ldots x_n$ is a $\Box$-theorem (178.1), it follows by Rule RN that $\Box(x_1 \ldots x_n F \to \Box x_1 \ldots x_n F)$. So by (174.2), $\mathcal{A} x_1 \ldots x_n F \equiv x_1 \ldots x_n F$. By the commutativity of $\equiv$, $x_1 \ldots x_n F \equiv \mathcal{A} x_1 \ldots x_n F$. ⋈

(180.1) Assume $O!x$. Then by definition of $O!$ (22.1) and Rule $=_{df}$E, $[\lambda x\, \Diamond E!x]x$. Now, independently, since $[\lambda x\, \Diamond E!x]\downarrow$, β-Conversion (48.2) yields:

(ϑ) $[\lambda x\, \Diamond E!x]x \equiv \Diamond E!x$

So it follows from what we've established thus far that $\Diamond E!x$. This implies, by the 5 axiom (45.3), that $\Box\Diamond E!x$. Note also that from the $\Box$-theorem (ϑ), its commuted form $\Diamond E!x \equiv [\lambda x\, \Diamond E!x]x$ is also as $\Box$-theorem. From this and $\Box\Diamond E!x$ it follows, by the Rule of Substitution (160.2), that $\Box[\lambda x\, \Diamond E!x]x$. Hence, by definition of $O!$ and Rule $=_{df}$I, $\Box O!x$. ⋈

(180.2) Assume $A!x$. Then by definition of $A!$ (22.2) and Rule $=_{df}$E, we know $[\lambda x\, \neg\Diamond E!x]x$. Independently, from the fact that $[\lambda x\, \neg\Diamond E!x]\downarrow$, it follows by β-Conversion that the following is a $\Box$-theorem:

(ϑ) $[\lambda x\, \neg\Diamond E!x]x \equiv \neg\Diamond E!x$

So it follows from what we've established that $\neg\Diamond E!x$. This implies $\Box\neg E!x$, by the modally strict equivalence (162.1). By the 4 schema (165.5) it follows that $\Box\Box\neg E!x$. From this and (162.1) again, we apply the Rule of Substitution (160.2) to obtain $\Box\neg\Diamond E!x$. This implies, by the commuted form of (ϑ) (which is also a $\Box$-theorem) and the Rule of Substitution (160.2) that $\Box[\lambda x\, \neg\Diamond E!x]x$. So by the definition of $A!$ and Rule $=_{df}$I, we have $\Box A!x$. ⋈

(180.3) From the $\Box$-theorem (180.1), it follows that $\Diamond O!x \to O!x$, by rule (166.2). $\bowtie$

(180.4) From the $\Box$-theorem (180.2) it follows that $\Diamond A!x \to A!x$, by rule (166.2). $\bowtie$

(180.5) ($\rightarrow$) By hypothetical syllogism from (180.3) and (180.1). ($\leftarrow$) By the T and T$\Diamond$ schemata. $\bowtie$

(180.6) ($\rightarrow$) By hypothetical syllogism from (180.4) and (180.2). ($\leftarrow$) By the T and T$\Diamond$ schemata. $\bowtie$

(180.7) Since $O!x \to \Box O!x$ (180.1) is a theorem, it follows by Rule RN that $\Box(O!x \to \Box O!x)$. So by (174.2), $\mathcal{A}O!x \equiv O!x$, which commutes to $O!x \equiv \mathcal{A}O!x$. $\bowtie$

(180.8) (Exercise)

(181) Let $\mu_1, \ldots, \mu_n$ ($n \geq 1$) be any distinct individual variables, and let φ be any formula in which $\nu_1, \ldots, \nu_n$ are any individual variables substitutable, respectively, for $\mu_1, \ldots, \mu_n$. We want to prove:[443]

$$(\xi)\ \ [\lambda\mu_1 \ldots \mu_n\, \varphi]{\downarrow} \to ([\lambda\mu_1 \ldots \mu_n\, \varphi]\nu_1 \ldots \nu_n \equiv \varphi^{\nu_1,\ldots,\nu_n}_{\mu_1,\ldots,\mu_n})$$

Note that by n applications of the Special Case of Rule $\forall$E (93.3), it suffices to prove a universal closure of (ξ), namely:

$$(\zeta)\ \ \forall\nu_1 \ldots \forall\nu_n([\lambda\mu_1 \ldots \mu_n\, \varphi]{\downarrow} \to ([\lambda\mu_1 \ldots \mu_n\, \varphi]\nu_1 \ldots \nu_n \equiv \varphi^{\nu_1,\ldots,\nu_n}_{\mu_1,\ldots,\mu_n}))$$

Note also that there may be free variables in φ not among $\nu_1, \ldots, \nu_n$ and $\mu_1, \ldots, \mu_n$. If there are such, then these variables occur free in (ζ). So suppose that $\alpha_1, \ldots, \alpha_m$ are the variables that occur free in (ζ). Then by m applications of the Special Case of Rule $\forall$E (93.3), it suffices to prove a universal closure of (ζ), namely (ω):

$$(\omega)\ \ \forall\alpha_1 \ldots \forall\alpha_m \forall\nu_1 \ldots \forall\nu_n([\lambda\mu_1 \ldots \mu_n\, \varphi]{\downarrow} \to ([\lambda\mu_1 \ldots \mu_n\, \varphi]\nu_1 \ldots \nu_n \equiv \varphi^{\nu_1,\ldots,\nu_n}_{\mu_1,\ldots,\mu_n}))$$

Clearly, there are no free variables in (ω). Now consider the specific list $x_1, \ldots, x_n$ of individual variables and perform the following transformation of φ:

- For each α_j ($1 \leq j \leq m$), if α_j is among the list of variables $x_1, \ldots, x_n$, let α'_j be a variable distinct from $x_1, \ldots, x_n$ and substitute α'_j for every free occurrence of α_j in φ and call the result φ''; otherwise, let α'_j be α_j. Thus, φ'' has no free occurrences of $x_1, \ldots, x_n$.

[443] I'm indebted to Uri Nodelman for helping me to work through a number of subtleties in the following proof; my original attempt had a number of gaps.

- Let φ' be an alphabetic variant of φ'' such that every bound occurrence of x_i in φ'' is replaced with a fresh variable.[444]

Then consider the following alphabetic variant of (ω):

$$(\omega')\ \ \forall\alpha'_1 \dots \forall\alpha'_m \forall\nu_1 \dots \forall\nu_n([\lambda\mu_1 \dots \mu_n\, \varphi']\downarrow \to ([\lambda\mu_1 \dots \mu_n\, \varphi']\nu_1 \dots \nu_n \equiv \varphi'^{\nu_1,\dots,\nu_n}_{\mu_1,\dots,\mu_n}))$$

To prove (ω), it suffices to prove (ω'), by the Rule of Alphabetic Variants. So if we can show that (ω') itself is an alphabetic variant of an axiom, we're done. Observe that there are no free variables in (ω') and, moreover, none of the variables $x_1,\dots,x_n$ occur free in any subformula or term in (ω') (since any free occurrence in φ would have been swapped out in the transformation from φ to φ''). Observe also that each variable x_i in the list $x_1,\dots,x_n$ is substitutable for μ_i in φ', for there are no operators binding x_i that could capture them after the transformation from φ'' to φ'. (One might recall here the discussion following theorem (99.13).) Consider the following alphabetic variant of (ω'):

$$\begin{aligned}&\forall\alpha'_1 \dots \forall\alpha'_m \forall x_1 \dots \forall x_n([\lambda x_1 \dots x_n\, (\varphi')^{x_1,\dots,x_n}_{\mu_1,\dots,\mu_n}]\downarrow \to\\ &\qquad ([\lambda x_1 \dots x_n\, (\varphi')^{x_1,\dots,x_n}_{\mu_1,\dots,\mu_n}]x_1 \dots x_n \equiv (\varphi')^{x_1,\dots,x_n}_{\mu_1,\dots,\mu_n}))\end{aligned}$$

But this claim is a universal closure of an instance of (48.2) and hence an axiom. ⋈

(183.1) For simplicity, we prove the theorem using the variables x_i and y_i as arbitrary stand-ins for any variables meeting the conditions of the theorem. (This will be easier to read than a proof using metavariables.) By hypothesis, for $1 \le i \le n$, the variable y_i is substitutable for the variable x_i in φ. Note that by n applications of GEN to an appropriate instance of Strengthened β-Conversion (181), we obtain:

$$\forall y_1 \dots \forall y_n([\lambda x_1 \dots x_n\, \varphi]\downarrow \to ([\lambda x_1 \dots x_n\, \varphi]y_1 \dots y_n \equiv \varphi^{y_1,\dots,y_n}_{x_1,\dots,x_n}))$$

Then by theorem (99.14), we may distribute the quantifiers $\forall y_1,\dots,\forall y_n$ over the conditional, so that we may conclude:

$$\forall y_1 \dots \forall y_n([\lambda x_1 \dots x_n\, \varphi]\downarrow) \to \forall y_1 \dots \forall y_n([\lambda x_1 \dots x_n\, \varphi]y_1 \dots y_n \equiv \varphi^{y_1,\dots,y_n}_{x_1,\dots,x_n}) \qquad ⋈$$

(183.2) For simplicity, we prove the theorem using the variables x_i and y_i as arbitrary stand-ins for any variables meeting the conditions of the theorem. (This will be easier to read than a proof using metavariables.) Assume:

$(\vartheta)\ \ [\lambda x_1 \dots x_n\, \varphi]\downarrow$

[444] The transformation from φ'' to φ' is done to ensure that every x_i is substitutable for μ_i. We want to avoid the situation in which the substitution of x_i for μ_i in φ would result in x_i's capture by a variable-binding operator binding x. But, by the transformation, φ' has no x_i-binding operators within it.

By hypothesis, $y_1, \ldots, y_n$ are substitutable, respectively, for $x_1, \ldots, x_n$ in φ and so $[\lambda x_1 \ldots x_n\, \varphi]$ gives rise to instances of (181):

$$[\lambda x_1 \ldots x_n\, \varphi]\downarrow \rightarrow ([\lambda x_1 \ldots x_n\, \varphi]y_1 \ldots y_n \equiv \varphi^{y_1,\ldots,y_n}_{x_1,\ldots,x_n})$$

So by GEN, we may infer:

$$\forall y_1 \ldots \forall y_n([\lambda x_1 \ldots x_n\, \varphi]\downarrow \rightarrow ([\lambda x_1 \ldots x_n\, \varphi]y_1 \ldots y_n \equiv \varphi^{y_1,\ldots,y_n}_{x_1,\ldots,x_n}))$$

But also, by hypothesis, none of $y_1, \ldots, y_n$ occur free in $[\lambda x_1 \ldots x_n\, \varphi]$. So from our last conclusion and (99.15), it follows that:

$$[\lambda x_1 \ldots x_n\, \varphi]\downarrow \rightarrow \forall y_1 \ldots \forall y_n([\lambda x_1 \ldots x_n\, \varphi]y_1 \ldots y_n \equiv \varphi^{y_1,\ldots,y_n}_{x_1,\ldots,x_n})$$

From this last result and our assumption (ϑ), it then follows that:

$$\forall y_1 \ldots \forall y_n([\lambda x_1 \ldots x_n\, \varphi]y_1 \ldots y_n \equiv \varphi^{y_1,\ldots,y_n}_{x_1,\ldots,x_n}) \qquad \bowtie$$

(184.1) We prove the (.a) and (b) forms of the rule together. Assume:

(ϑ) $[\lambda \mu_1 \ldots \mu_n\, \varphi]\kappa_1 \ldots \kappa_n$

Then by axiom (39.5.a), we know all of the following:

(A) $[\lambda \mu_1 \ldots \mu_n\, \varphi]\downarrow$

(B_1) $\kappa_1\downarrow$

$\vdots$

(B_n) $\kappa_n\downarrow$

By hypothesis, $[\lambda \mu_1 \ldots \mu_n\, \varphi]$ has no free variables. Now let $\nu_1, \ldots, \nu_n$ be any individual variables substitutable, respectively, for $\mu_1, \ldots, \mu_n$ in φ. Since $\nu_1, \ldots, \nu_n$ aren't free in $[\lambda \mu_1 \ldots \mu_n\, \varphi]$, it follows from (A) by (183.2) that:

$$\forall \nu_1 \ldots \forall \nu_n([\lambda \mu_1 \ldots \mu_n\, \varphi]\nu_1 \ldots \nu_n \equiv \varphi^{\nu_1,\ldots,\nu_n}_{\mu_1,\ldots,\mu_n})$$

But from (B_1) – (B_n) and the fact that $\kappa_1, \ldots, \kappa_n$ are substitutable, respectively, for $\mu_1, \ldots, \mu_n$ in φ, it follows by n applications of Rule $\forall$E that:

(ξ) $[\lambda \mu_1 \ldots \mu_n\, \varphi]\kappa_1 \ldots \kappa_n \equiv (\varphi^{\nu_1,\ldots,\nu_n}_{\mu_1,\ldots,\mu_n})^{\kappa_1,\ldots,\kappa_n}_{\nu_1,\ldots,\nu_n}$

Since, for all i such that $1 \leq i \leq n$, both κ_i and ν_i are substitutable for μ_i in φ, we can establish independently that $(\varphi^{\nu_1,\ldots,\nu_n}_{\mu_1,\ldots,\mu_n})^{\kappa_1,\ldots,\kappa_n}_{\nu_1,\ldots,\nu_n}$ just is $\varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n}$ (exercise). Hence, (ξ) becomes $[\lambda \mu_1 \ldots \mu_n\, \varphi]\kappa_1 \ldots \kappa_n \equiv \varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n}$. But then, given ($\vartheta$), it follows that $\varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n}$. So, by conditional proof, we've established:

(ζ) $\vdash [\lambda \mu_1 \ldots \mu_n\, \varphi]\kappa_1 \ldots \kappa_n \rightarrow \varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n}$

So by (63.10), the (.a) form of our rule follows from (ζ):

(.a) $[\lambda\mu_1 \ldots \mu_n\, \varphi]\kappa_1 \ldots \kappa_n \vdash \varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n}$

Moreover, from (ζ) and the rules of contraposition (80), it follows that:

$$\vdash \neg\varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n} \rightarrow \neg[\lambda\mu_1 \ldots \mu_n\, \varphi]\kappa_1 \ldots \kappa_n$$

But this implies, by (63.10), the (.b) form of our rule:

(.b) $\neg\varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n} \vdash \neg[\lambda\mu_1 \ldots \mu_n\, \varphi]\kappa_1 \ldots \kappa_n$ ⋈

(184.2) We prove only the (.a) form of the rule and leave the (.b) form as an exercise. Assume all of the following:

(A) $[\lambda\mu_1 \ldots \mu_n\, \varphi]\downarrow$

(B_1) $\kappa_1\downarrow$

$\vdots$

(B_n) $\kappa_n\downarrow$

(C) $\varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n}$

By hypothesis, $[\lambda\mu_1 \ldots \mu_n\, \varphi]$ has no free variables. Now let $\nu_1,\ldots,\nu_n$ be any individual variables substitutable, respectively, for $\mu_1,\ldots,\mu_n$ in φ. Since $\nu_1,\ldots,\nu_n$ aren't free in $[\lambda\mu_1 \ldots \mu_n\, \varphi]$, it follows from (A) by (183.2) that:

$$\forall\nu_1 \ldots \forall\nu_n([\lambda\mu_1 \ldots \mu_n\, \varphi]\nu_1 \ldots \nu_n \equiv \varphi^{\nu_1,\ldots,\nu_n}_{\mu_1,\ldots,\mu_n})$$

But from (B_1) – (B_n) and the fact that $\kappa_1,\ldots,\kappa_n$ are substitutable, respectively, for $\mu_1,\ldots,\mu_n$ in φ, it follows by n applications of Rule ∀E that:

(ϑ) $[\lambda\mu_1 \ldots \mu_n\, \varphi]\kappa_1 \ldots \kappa_n \equiv (\varphi^{\nu_1,\ldots,\nu_n}_{\mu_1,\ldots,\mu_n})^{\kappa_1,\ldots,\kappa_n}_{\nu_1,\ldots,\nu_n}$

Since, for all i such that $1 \leq i \leq n$, both κ_i and ν_i are substitutable for μ_i in φ, we can establish independently that $(\varphi^{\nu_1,\ldots,\nu_n}_{\mu_1,\ldots,\mu_n})^{\kappa_1,\ldots,\kappa_n}_{\nu_1,\ldots,\nu_n}$ just is $\varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n}$ (exercise). Hence, it follows from (ϑ) that $[\lambda\mu_1 \ldots \mu_n\, \varphi]\kappa_1 \ldots \kappa_n \equiv \varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n}$. But then, by (C), it follows that $[\lambda\mu_1 \ldots \mu_n\, \varphi]\kappa_1 \ldots \kappa_n$. Hence, by $n+2$ applications of conditional proof, we have established that:

$$\vdash [\lambda\mu_1 \ldots \mu_n\varphi]\downarrow \rightarrow (\kappa_1\downarrow \rightarrow (\ldots \rightarrow (\kappa_n\downarrow \rightarrow (\varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n} \rightarrow [\lambda\mu_1 \ldots \mu_n\varphi]\kappa_1 \ldots \kappa_n))))$$

So by $n+1$ applications of (63.10), it follows that

$$[\lambda\mu_1 \ldots \mu_n\, \varphi]\downarrow, \kappa_1\downarrow, \ldots, \kappa_n\downarrow, \varphi^{\kappa_1,\ldots,\kappa_n}_{\mu_1,\ldots,\mu_n} \vdash [\lambda\mu_1 \ldots \mu_n\, \varphi]\kappa_1 \ldots \kappa_n$$ ⋈

(186.1) Let Π^n be any n-ary relation term ($n \geq 0$) in which none of $x_1,\ldots,x_n$ occur free. Now assume $\Pi^n{\downarrow}$. Independently, we know that the closures of η-Conversion (48.3) are axioms. Hence, we know, for every $n \geq 0$:

(ξ) $\forall F^n([\lambda x_1 \ldots x_n\, F^n x_1 \ldots x_n] = F^n)$

By hypothesis, $x_1,\ldots,x_n$ aren't free in Π^n, and so Π^n is substitutable for F in $[\lambda x_1 \ldots x_n\, F^n x_1 \ldots x_n] = F^n$. Hence, the provisos for applying Rule $\forall$E (93.1) are met so that we may conclude $[\lambda x_1 \ldots x_n\, \Pi^n x_1 \ldots x_n] = \Pi^n$. ⋈

(186.2) Assume $\Pi^n{\downarrow}$, where $x_1,\ldots,x_n$ are any distinct individual variables none of which occur free in Π^n. Then by (186.1), it follows that:

(ϑ) $[\lambda x_1 \ldots x_n\, \Pi^n x_1 \ldots x_n] = \Pi^n$

But then, where $\nu_1,\ldots,\nu_n$ are any distinct variables not free in Π^n, we know:

$[\lambda \nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$ is an alphabetic variant of $[\lambda x_1 \ldots x_n\, \Pi^n x_1 \ldots x_n]$

Hence $[\lambda \nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n] = \Pi^n$ is an alphabetic variant of (ϑ) and so follows from (ϑ) by the Rule of Alphabetic Variants (114). ⋈

(187) By hypothesis:

(a) $\vdash \rho{\downarrow}$,

(b) ρ' is an η-variant of ρ witnessed by the sequence $\rho_1,\ldots,\rho_m$, for which $\rho = \rho_1$ and $\rho' = \rho_m$ ($m \geq 1$), and

(c) $\vdash \Pi^n{\downarrow}$ ($n \geq 0$) whenever ρ_{i+1} is an immediate η-variant of ρ_i with respect to Π^n, for each i such that $1 \leq i \leq m-1$.

We want to show $\vdash \rho = \rho'$. We prove this by cases.

Case 1: ρ' is an *immediate* η-variant of ρ with respect to Π^n. Then:

- by definition (185.4), either (i) Π^n is a subterm of ρ and ρ' results from ρ by replacing Π^n by an η-expansion $[\lambda \nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$ or (ii) the elementary expression $[\lambda \nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$ is a subterm of ρ and ρ' results from ρ by replacing $[\lambda \nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$ by its η-contraction Π^n, and
- by definition (185.5), the sequence $\langle \rho, \rho' \rangle$ is a witness to the fact that ρ' is an η-variant of ρ.

By assumption (c), i.e., $\vdash \Pi^n{\downarrow}$, and so from the fact that (186.2) is a theorem, it follows by (63.6) that:

(ϑ) $\vdash [\lambda \nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n] = \Pi^n$

By symmetry of identity, we also know:

(ξ) $\vdash \Pi^n = [\lambda \nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$

Moreover, since $\vdash \rho\downarrow$ by assumption (a), we know by Rule =I that:

(ζ) $\vdash \rho = \rho$

Then in case (i), (ξ) and (ζ) imply $\rho = \rho'$ by Rule =E (110), and in case (ii), (ϑ) and (ζ) imply $\rho = \rho'$ by Rule =E.

Case 2: ρ' is an η-variant of ρ but not an immediate η-variant with respect to Π^n. Since ρ' is not an immediate η-variant of ρ, it follows from assumption (b) that $m \geq 3$. By assumption (c), we know that for $1 \leq i \leq m-1$, if ρ_{i+1} is an immediate η-variant of ρ_i with respect to the n-ary relation term Π^n ($n \geq 1$), then $\vdash \Pi^n\downarrow$. So for each i, $1 \leq i \leq m-1$, it follows from *Case 1* that $\vdash \rho_i = \rho_{i+1}$. Hence, by $m-2$ applications of the transitivity of identity (117.3), it follows that $\vdash \rho = \rho'$ ⋈

(188.1) Assume both:

(A) $[\lambda z_1 \ldots z_n\, \chi_y^{\imath x\varphi}]\downarrow$

(B) $\imath x\varphi = \imath x\psi$

Now (A) implies, by Rule =I:

(C) $[\lambda z_1 \ldots z_n\, \chi_y^{\imath x\varphi}] = [\lambda z_1 \ldots z_n\, \chi_y^{\imath x\varphi}]$

But where (ϑ) is:

(ϑ) $[\lambda z_1 \ldots z_n\, \chi] = [\lambda z_1 \ldots z_n\, \chi]$

then (C) is, by the definition of substitutions, $(\vartheta)_y^{\imath x\varphi}$. Now since $\imath x\varphi$ and $\imath x\psi$ are, by hypothesis, both substitutable for y in $[\lambda z_1 \ldots z_n\, \chi]$, they are both substitutable for y in (ϑ). Hence, by Rule =E (110), $(\vartheta)_y^{\imath x\varphi}$ and (B), we may infer (ϑ'), i.e., the result of substituting $\imath x\psi$ for zero or more occurrences of $\imath x\varphi$ in $(\vartheta)_y^{\imath x\varphi}$. But the following is such a (ϑ'):

$$[\lambda z_1 \ldots z_n\, \chi_y^{\imath x\varphi}] = [\lambda z_1 \ldots z_n\, \chi']$$

since, by hypothesis, χ' is the result of substituting $\imath x\psi$ for one or more occurrences of $\imath x\varphi$ in $\chi_y^{\imath x\varphi}$. ⋈

(188.2) (Exercise)

(189) ($\rightarrow$) Exercise. ($\leftarrow$) For conditional proof, assume:

(ϑ) $\forall x(xF^1 \equiv xG^1)$

to show $F^1 = G^1$. By theorem (116.1), we have to show:

$$\Box\forall x(xF^1 \equiv xG^1)$$

We can't simply apply RN (68) to (ϑ), for that would violate the conditions of application of the rule. But by the Barcan Formula (167.1), it suffices to show:

$$\forall x\Box(xF^1 \equiv xG^1)$$

Since x is not free in our assumption (ϑ), it suffices by GEN to show $\Box(xF^1 \equiv xG^1)$. We do this as follows. From (ϑ), it follows by Rule $\forall$E (93.3) that:

(ζ) $xF^1 \equiv xG^1$

Note that we independently know, by (179.2), both:

(ξ) $xF^1 \equiv \Box xF^1$

(ω) $xG^1 \equiv \Box xG^1$

So we can show $\Box xF^1 \equiv \Box xG^1$ as follows:

$$\begin{array}{llll} \Box xF^1 & \equiv & xF^1 & \text{by } (\xi)\text{, commutativity of } \equiv \\ & \equiv & xG^1 & \text{by } (\zeta) \\ & \equiv & \Box xG^1 & \text{by } (\omega) \end{array}$$

From this, we can reach $\Box(xF^1 \equiv xG^1)$, by the right-to-left direction of (179.5). ⋈

(191.1) Suppose φ contains (a) no free occurrences of $x_1,\ldots,x_n$ in encoding position (9.1), and (b) no free occurrences of F^n. Then, from (a) and (39.2), we know $[\lambda x_1 \ldots x_n\, \varphi]{\downarrow}$. It then follows by β-Conversion (48.2) that:

$$[\lambda x_1 \ldots x_n\, \varphi]x_1 \ldots x_n \equiv \varphi$$

So by n applications of GEN:

$$\forall x_1 \ldots \forall x_n([\lambda x_1 \ldots x_n\, \varphi]x_1 \ldots x_n \equiv \varphi)$$

Since this claim has been established by a modally strict proof, it follows by RN (68) that:

(ϑ) $\Box\forall x_1 \ldots \forall x_n([\lambda x_1 \ldots x_n\, \varphi]x_1 \ldots x_n \equiv \varphi)$

Since F doesn't occur free in φ, the previous line has the form $\psi_F^{[\lambda x_1 \ldots x_n\, \varphi]}$ when ψ is the formula:

$$\Box\forall x_1 \ldots \forall x_n(Fx_1 \ldots x_n \equiv \varphi)$$

From (ϑ) and the fact that $[\lambda x_1 \ldots x_n\, \varphi]{\downarrow}$, it follows by $\exists$I that $\exists F\psi$, i.e., that:[445]

[445] Thus, the fact that F doesn't occur free in φ is essential to the proof, for otherwise the free occurrence of F in φ would be captured if we were to introduce $\exists F$ to existentially generalize on the relation term $[\lambda x_1 \ldots x_n\, \varphi]$. This would be an invalid application of $\exists$I, like the ones discussed in the penultimate paragraph of (101).

$$\exists F^n \Box \forall x_1 \ldots \forall x_n (F^n x_1 \ldots x_n \equiv \varphi) \qquad \bowtie$$

(191.2) This is just the unary case of (191.1). $\bowtie$

(192.1) Suppose, for reductio:

$$[\lambda x\, \exists G(xG \,\&\, \neg Gx)]\downarrow$$

Note that the above λ-expression has no free variables. Now consider the following instance of the Comprehension Principle for Abstract Objects (53):

$$\exists x(A!x \,\&\, \forall F(xF \equiv F = [\lambda x\, \exists G(xG \,\&\, \neg Gx)]))$$

Suppose a is such an object, so that we know $a\downarrow$ and:

$$(\xi)\ \ A!a \,\&\, \forall F(aF \equiv F = [\lambda x\, \exists G(xG \,\&\, \neg Gx)])$$

Now, for notational convenience, abbreviate $[\lambda x \exists G(xG \& \neg Gx)]$ as $[\lambda x \varphi]$. Since either $[\lambda x \varphi]a$ or $\neg[\lambda x\, \varphi]a$, we may reason each case to a contradiction. Assume $[\lambda x \varphi]a$. Then by Rule $\overrightarrow{\beta}$C (184.1.a), $\exists G(aG \& \neg Ga)$. Let P be such a property, so that we know $P\downarrow$, aP, and $\neg Pa$. But from the first two, it follows from the second conjunct of (ξ) by $\forall$E that $P = [\lambda x\, \varphi]$. So $\neg[\lambda x\, \varphi]a$, contrary to hypothesis. Then assume $\neg[\lambda x\, \varphi]a$. From this and the facts that $[\lambda x\, \varphi]\downarrow$ and $a\downarrow$, Rule $\overleftarrow{\beta}$C (184.2.b) implies $\neg\exists G(aG \& \neg Ga)$, i.e., $\forall G(aG \to Ga)$. Hence, $a[\lambda x \varphi] \to [\lambda x \varphi]a$. But since $[\lambda x\, \varphi]\downarrow$, the second conjunct of (ξ) implies $a[\lambda x\, \varphi] \equiv [\lambda x\, \varphi] = [\lambda x\, \varphi]$ and Rule =I (118) implies $[\lambda x\, \varphi] = [\lambda x\, \varphi]$. Hence, $a[\lambda x\, \varphi]$, and so $[\lambda x\, \varphi]a$. Contradiction. $\bowtie$

(192.2) (Exercise)

(192.3) Assume, for reductio:

$$(\vartheta)\ \ \forall y([\lambda z\, z = y]\downarrow)$$

Now by the Comprehension Principle for Abstract Objects (53), we know:

$$\exists x(A!x \,\&\, \forall F(xF \equiv \exists y(F = [\lambda z\, z = y] \,\&\, \neg yF)))$$

Let a be such an object, so that we know both $a\downarrow$ and:

$$(\xi)\ \ A!a \,\&\, \forall F(aF \equiv \exists y(F = [\lambda z\, z = y] \,\&\, \neg yF))$$

Then it follows from (ϑ) that $[\lambda z\, z = a]\downarrow$. So we may instantiate $[\lambda z\, z = a]$ into the second conjunct of (ξ) to conclude:

$$(\zeta)\ \ a[\lambda z\, z = a] \equiv \exists y([\lambda z\, z = a] = [\lambda z\, z = y] \,\&\, \neg y[\lambda z\, z = a])$$

Now either $a[\lambda z\, z = a]$ or $\neg a[\lambda z\, z = a]$, but both cases lead to contradiction.

- Assume $a[\lambda z\, z = a]$. Then by (ζ):

$$\exists y([\lambda z\, z=a]=[\lambda z\, z=y]\,\&\,\neg y[\lambda z\, z=a])$$

Suppose b is such an object, so that $[\lambda z\, z=a]=[\lambda z\, z=b]\,\&\,\neg b[\lambda z\, z=a]$. Since $a\downarrow$, we know by Rule =I that $a=a$. From this and the facts that $[\lambda z\, z=a]\downarrow$ and $a\downarrow$, it follows by Rule $\overleftarrow{\beta}$C (184.2.a) that $[\lambda z\, z=a]a$. But we know $[\lambda z z=a]=[\lambda z z=b]$ and so by Rule =E, $[\lambda z z=b]a$. It follows by Rule $\overrightarrow{\beta}$C (184.1.a) that $a=b$, i.e., $b=a$. But since we established $\neg b[\lambda z\, z=a]$, it follows that $\neg a[\lambda z\, z=a]$, contrary to hypothesis.

- Assume $\neg a[\lambda z\, z=a]$. Then by (ζ):

$$\neg\exists y([\lambda z\, z=a]=[\lambda z\, z=y]\,\&\,\neg y[\lambda z\, z=a])$$

i.e.,

$$\forall y([\lambda z\, z=a]=[\lambda z\, z=y]\rightarrow y[\lambda z\, z=a])$$

Since $a\downarrow$, we may instantiate the above to a to obtain:

$$[\lambda z\, z=a]=[\lambda z\, z=a]\rightarrow a[\lambda z\, z=a]$$

But since we've established that $[\lambda z\, z=a]\downarrow$, we know by Rule =I that the antecedent of this last result holds. Hence, $a[\lambda z\, z=a]$, contrary to hypothesis. ⋈

(192.4) (Exercise) [Note: it is easier to prove this theorem with the help of a later theorem, namely, (269). But the preceding theorem shows the way towards a proof that doesn't cite (269).]

(192.5) (Exercise)

(193.1)★ Assume $\forall xGx$. To prove our theorem, we need the following lemma, which holds under this assumption:

Lemma: $\forall x(G\imath y(y=x\,\&\,\exists H(xH\,\&\,\neg Hx))\equiv\exists H(xH\,\&\,\neg Hx))$

Let's suppose, for the moment, that the Lemma holds (we'll prove it at the end). Then assume, for reductio:

$$[\lambda x\, G\imath y(y=x\,\&\,\exists H(xH\,\&\,\neg Hx))]\downarrow$$

Now suppose we simply abbreviate the λ-expression in the above claim as $[\lambda x\,\varphi]$. Then since $[\lambda x\,\varphi]$ has no free variables (and, in particular, x doesn't occur free in $[\lambda x\,\varphi]$) and x is substitutable for itself in the matrix of $[\lambda x\,\varphi]$, it follows from the above claim, by (183.2), that:

(ϑ) $\forall x([\lambda x\,\varphi]x\equiv G\imath y(y=x\,\&\,\exists H(xH\,\&\,\neg Hx)))$

But from this and the Lemma, it follows by (99.10) that:

$$\forall x([\lambda x\, \varphi]x \equiv \exists H(xH \,\&\, \neg Hx))$$

Since $[\lambda x\, \varphi]\!\downarrow$, it follow by $\exists$I that:

$$\exists F \forall x(Fx \equiv \exists H(xH \,\&\, \neg Hx))$$

But this contradicts (192.2). So it remains only to show the Lemma holds under the assumption $\forall xGx$.

Proof of the Lemma: We gave a proof sketch of this in footnote 29. But here is a proof sketch that cites the theorems we've now established. By GEN, it suffices to show:

$$G\imath y(y = x \,\&\, \exists H(xH \,\&\, \neg Hx)) \equiv \exists H(xH \,\&\, \neg Hx)$$

But this follows, by biconditional syllogism, from the following two biconditionals:

(A) $G\imath y(y = x \,\&\, \exists H(xH \,\&\, \neg Hx)) \equiv \exists! y(y = x \,\&\, \exists H(xH \,\&\, \neg Hx))$

(B) $\exists! y(y = x \,\&\, \exists H(xH \,\&\, \neg Hx)) \equiv \exists H(xH \,\&\, \neg Hx)$

Proof of (A). We prove both directions. ($\rightarrow$) Assume $G\imath y(y = x \& \exists H(xH \& \neg Hx))$. Then by axiom (39.5.a), $\imath y(y = x \,\&\, \exists H(xH \,\&\, \neg Hx))\!\downarrow$. So by (144.1)$\bigstar$, $\exists! y(y = x \,\&\, \exists H(xH \,\&\, \neg Hx))$. ($\leftarrow$) Assume $\exists! y(y = x \,\&\, \exists H(xH \,\&\, \neg Hx))$. Let a be such an object so that, by the definition of the unique existence quantifier, we know both:

$$a = x \,\&\, \exists H(xH \,\&\, \neg Hx))$$

$$\forall z(z = x \,\&\, \exists H(xH \,\&\, \neg Hx) \rightarrow z = a)$$

But recall that we're under the assumption $\forall xGx$. Hence Ga. So if we conjoin the above two claims with Ga, it follows by $\exists$I that there is an object y (namely a) such that (i) $y = x \& \exists H(xH \& \neg Hx)$, (ii) $\forall z(z = x \& \exists H(xH \& \neg Hx) \rightarrow z = y)$, and (iii) Gy. Hence, by Russell's theory of descriptions (143)$\bigstar$, $G\imath y(y = x \& \exists H(xH \& \neg Hx))$.

Proof of (B). Again, we prove both directions. ($\rightarrow$) Assume $\exists! y(y = x \& \exists H(xH \& \neg Hx))$. Then, as before, let a be such an object, so that we know, among other things, that $a = x \,\&\, \exists H(xH \,\&\, \neg Hx))$. Then it follows that $\exists H(xH \,\&\, \neg Hx)$. ($\leftarrow$) Assume $\exists H(xH \,\&\, \neg Hx)$. Since we know $x = x$ by the reflexivity of identity, it follows that:

(ζ) $x = x \,\&\, \exists H(xH \,\&\, \neg Hx)$

Moreover, it is trivial that:

(ω) $\forall z(z = x \,\&\, \exists H(xH \,\&\, \neg Hx) \rightarrow z = x)$

So if we conjoin (ζ) and (ω) and apply $\exists$I, it follows that:

$$\exists y(y=x \,\&\, \exists H(xH \,\&\, \neg Hx) \,\&\, \forall z(z=x \,\&\, \exists H(xH \,\&\, \neg Hx) \rightarrow z=y))$$

So by the definition of the unique existence quantifier, $\exists!y(y=x \,\&\, \exists H(xH \,\&\, \neg Hx))$. ⋈

(193.2)★ (Exercise)

(194) Let φ be any formula with no free occurrences of p. Since $\varphi \equiv \varphi$ is an instance of the modally strict theorem (88.3.a), it follows by RN that $\Box(\varphi \equiv \varphi)$. But by (104.2), we know that $\varphi\!\downarrow$. Since p doesn't occur free in φ, we may use Rule $\exists$I (101.1) to conclude $\exists p\Box(p \equiv \varphi)$. ⋈

(195.1) Assume $\Diamond\neg\forall x_1 \ldots \forall x_n(F^n x_1 \ldots x_n \equiv G^n x_1 \ldots x_n)$, which by (158.11), implies $\neg\Box\forall x_1 \ldots \forall x_n(F^n x_1 \ldots x_n \equiv G^n x_1 \ldots x_n)$. Then by modus tollens and (108.1), $F^n \neq G^n$. ⋈

(195.2) [The following proof was found by Daniel Kirchner's implementation in Isabelle and so replaces the longer proof I had originally constructed; see Kirchner 2017 [2021] and 2022.] Suppose G is substitutable for F in φ and φ' is the result of substituting G for one or more free occurrences of F in φ. Our proof strategy is:

(A) Prove, by modally strict means, that $(\varphi \not\equiv \varphi') \rightarrow F \neq G$.

(B) By RM$\Diamond$, conclude $\Diamond(\varphi \not\equiv \varphi') \rightarrow \Diamond F \neq G$.

(C) But by (170.3), we know $\Diamond F \neq G \rightarrow F \neq G$.

(D) From (B) and (C) it follows that $\Diamond(\varphi \not\equiv \varphi') \rightarrow F \neq G$.

Since (B) – (D) are straightforward, it remains to show (A). So assume $\varphi \not\equiv \varphi'$. Hence by (88.4.h), $(\varphi \,\&\, \neg\varphi') \vee (\neg\varphi \,\&\, \varphi')$. We then show that $F \neq G$ in both cases.

- $\varphi \,\&\, \neg\varphi'$. So $\neg(\varphi \rightarrow \varphi')$. But since G is substitutable for F in φ and φ' is the result of substituting G for one or more occurrences of F in φ, we know $F=G \rightarrow (\varphi \rightarrow \varphi')$ is an instance of the axiom for the substitution of identicals (41). Hence, $F \neq G$.

- $\neg\varphi \,\&\, \varphi'$. Then since $\varphi' \,\&\, \neg\varphi$, we know $\neg(\varphi' \rightarrow \varphi)$. But note that if G is substitutable for F in φ and φ' is the result of substituting G for one or more free occurrences of F in φ, then F is substitutable for G in φ' and φ is the result of substituting F for free occurrences of G in φ'. So as an instance (41) we know $G=F \rightarrow (\varphi' \rightarrow \varphi)$. Hence $G \neq F$, and so $F \neq G$, by the symmetry of identity and modus tollens. ⋈

(195.3) – (195.4) (Exercises)

(197.1) Let Π^n be any n-ary relation term in which $x_1,\ldots,x_n$ don't occur free ($n \geq 0$). Then $[\lambda x_1 \ldots x_n \neg \Pi^n x_1 \ldots x_n]$ is a core λ-expression and so as an instance of axiom (39.2) we know $[\lambda x_1 \ldots x_n \neg \Pi^n x_1 \ldots x_n]\downarrow$. ⋈

(197.2) Let Π be any n-ary relation term in which $x_1,\ldots,x_n$ don't occur free ($n \geq 0$). Note that the Rule of Definition by Identity asserts the following with respect to the instance of definition (196) in which Π is substituted for F:

$$(\vartheta)\quad ([\lambda x_1 \ldots x_n \neg \Pi x_1 \ldots x_n]\downarrow \rightarrow (\overline{\Pi} = [\lambda x_1 \ldots x_n \neg \Pi x_1 \ldots x_n]))\ \&$$
$$(\neg[\lambda x_1 \ldots x_n \neg \Pi x_1 \ldots x_n]\downarrow \rightarrow \neg\overline{\Pi}\downarrow)$$

But by (197.1), we know $[\lambda x_1 \ldots x_n \neg \Pi x_1 \ldots x_n]\downarrow$. Hence, we may infer from the first conjunct of (ϑ) that $\overline{\Pi} = [\lambda x_1 \ldots x_n \neg \Pi x_1 \ldots x_n]$. ⋈

(197.3) (Exercise)

(199.1) We reason by cases: *Case*: $n \geq 1$ and *Case*: $n = 0$.

Case $n \geq 1$. By (197.2), we know $\overline{F}^n = [\lambda x_1 \ldots x_n \neg F^n x_1 \ldots x_n]$. By the symmetry of identity, this implies $[\lambda x_1 \ldots x_n \neg F^n x_1 \ldots x_n] = \overline{F}^n$. It follows from these facts, respectively, by the axiom for the substitution of identicals (41) that:

$$\overline{F}^n x_1 \ldots x_n \rightarrow [\lambda x_1 \ldots x_n \neg F^n x_1 \ldots x_n]x_1 \ldots x_n$$

$$[\lambda x_1 \ldots x_n \neg F^n x_1 \ldots x_n]x_1 \ldots x_n \rightarrow \overline{F}^n x_1 \ldots x_n$$

Hence:

$$(\vartheta)\quad \overline{F}^n x_1 \ldots x_n \equiv [\lambda x_1 \ldots x_n \neg F^n x_1 \ldots x_n]x_1 \ldots x_n$$

Moreover, since we know $[\lambda x_1 \ldots x_n \neg F^n x_1 \ldots x_n]\downarrow$, it is a consequence of β-Conversion (48.2) that:

$$(\xi)\quad [\lambda x_1 \ldots x_n \neg F^n x_1 \ldots x_n]x_1 \ldots x_n \equiv \neg F^n x_1 \ldots x_n$$

But (ϑ) and (ξ) imply:

$$\overline{F}^n x_1 \ldots x_n \equiv \neg F^n x_1 \ldots x_n$$

Case $n = 0$. By (197.2), we know $\overline{p} = [\lambda\, \neg p]$. By theorem (108.2), this implies $\Box(\overline{p} \equiv [\lambda\, \neg p])$, and so $\overline{p} \equiv [\lambda\, \neg p]$, by the T schema. But $[\lambda\, \neg p] \equiv \neg p$, by (111.2). Hence, $\overline{p} \equiv \neg p$. ⋈

(199.2) (Exercise)

(199.3) This is the 0-ary case of (199.1). ⋈

(199.4) This is the 0-ary case of (199.2). ⋈

(199.5) Assume, for reductio, that $F^n = \overline{F}^n$. Then by (107.1) and (107.2), both $F^n\downarrow$ and $\overline{F}^n\downarrow$, and so we can apply the symmetry of identity (117.2) to conclude $\overline{F}^n = F^n$. Since theorem (199.1) is that $\overline{F}^n x_1 \ldots x_n \equiv \neg F^n x_1 \ldots x_n$, it follows by Rule =E that $F^n x_1 \ldots x_n \equiv \neg F^n x_1 \ldots x_n$, which is a contradiction (88.3.c). ⋈

(199.6) This is the 0-ary case of (199.5). ⋈

(199.7) By definition (196), we know $\overline{p} = [\lambda\, \neg p]$. Moreover, by theorem (111.1), we also know $[\lambda\, \neg p] = \neg p$. Clearly, all of the terms mentioned thus far are significant, and so by the transitivity of identity (117.3), $\overline{p} = \neg p$. ⋈

(199.8) Assume $p = q$. Although it is a theorem that every formula is significant and, hence, that $(\neg p)\downarrow$, we prove the latter without appeal to this theorem, for when we get to the type-theoretic version of object theory, there will be formulas that aren't significant and so a proof based on the theorem that every formula is significant wouldn't transfer to type theory. So note that $(\neg p)\downarrow$ follows directly from two principles: definition (20.3), which has the instance $(\neg p)\downarrow \equiv [\lambda x\, \neg p]\downarrow$, and axiom (39.2), which has $[\lambda x\, \neg p]\downarrow$ as an instance. Since $(\neg p)\downarrow$, it follows by Rule =I (118.1) that $\neg p = \neg p$. Hence $\neg p = \neg q$, by Rule =E (110). ⋈

(199.9) (Exercise)

(202.1) ($\rightarrow$) Assume *NonContingent*(F). Then by the unary case of definition (200.3) and Rule $\equiv_{df}$E (90.2), we know:

(A) $\mathit{Necessary}(F) \vee \mathit{Impossible}(F)$

Given (A) and definition (200.1), we may infer, by a relevant instance of (88.8.h) and biconditional syllogism, that:

(B) $\Box\forall x Fx \vee \mathit{Impossible}(F)$

Independently, by Rule $\equiv$S of Biconditional Simplication (91.1), we can infer, from definition (200.2) and the fact that $F\downarrow$ (recall Remark (201)), that:

(C) $\mathit{Impossible}(F) \equiv \Box\forall x \neg Fx$

So from (B) and (C), it follows by (88.8.g) and biconditional syllogism that:

(D) $\Box\forall x Fx \vee \Box\forall x \neg Fx$

Now we know by (199.2) and (199.1), respectively, that the following are $\Box$-theorems: $\neg\overline{F}x \equiv Fx$ and $\overline{F}x \equiv \neg Fx$. Applying the commutativity of $\equiv$ to each, we therefore have the following $\Box$-theorems:

(E) $Fx \equiv \neg\overline{F}x$

(F) $\neg Fx \equiv \overline{F}x$

Hence, from (D), (E), and (F), it follows by two applications of the Rule of Substitution (160.2) that:

$$\Box\forall x\neg\overline{F}x \vee \Box\forall x\overline{F}x$$

And by the commutativity of $\vee$, this implies:

(G) $\Box\forall x\overline{F}x \vee \Box\forall x\neg\overline{F}x$

From the commuted form of definition (200.1) and (G) it follows by an instance of (88.8.h) that:

(H) $Necessary(\overline{F}) \vee \Box\forall x\neg\overline{F}x$

Independently, by Rule $\equiv$S of Biconditional Simplification (91.1), we know that definition (200.2) and the fact that $\overline{F}\downarrow$ jointly imply the following equivalence:

(I) $Impossible(\overline{F}) \equiv \Box\forall x\neg\overline{F}x$

If we start with the commuted form of (I) and then consider (H), we may conclude, by (88.8.g), that:

$$Necessary(\overline{F}) \vee Impossible(\overline{F})$$

Hence, $NonContingent(\overline{F})$, by definition (200.3). ($\leftarrow$) Reverse the reasoning. ⋈

(202.2) ($\rightarrow$) Assume $Contingent(F)$. Its definition (200.4) and the axiom $F\downarrow$ imply, by Rule $\equiv$S of Biconditional Simplification (91.1):

(A) $\neg(Necessary(F) \vee Impossible(F))$

From (A) and the modally strict biconditional introduced by definition (200.1), we may infer, by the Rule of Substitution for Defined Formulas (160.3), that:

(B) $\neg(\Box\forall xFx \vee Impossible(F))$

Independently, by Rule $\equiv$S of Biconditional Simplification (91.1), definition (200.2) and the axiom $F\downarrow$ imply the modally strict theorem:

(C) $Impossible(F) \equiv \Box\forall x\neg Fx$

So by the Rule of Substitution (160.2), it follows from (B) and (C) that:

$$\neg(\Box\forall xFx \vee \Box\forall x\neg Fx)$$

By De Morgan's Law (88.5.d), it follows that:

$$\neg\Box\forall xFx \ \&\ \neg\Box\forall x\neg Fx$$

Using (158.11) on both conjuncts, it follows that:

$$\Diamond\neg\forall xFx \ \& \ \Diamond\neg\forall x\neg Fx$$

If we consider the form of the left conjunct and consider a modally-strict quantifier-negation $\Box$-theorem (103.2), we may infer by the Rule of Substitution (160.2) that:

$$\Diamond\exists x\neg Fx \ \& \ \Diamond\neg\forall x\neg Fx$$

If we consider the form of the right conjunct and consider the modally strict biconditional introduced by the definition of $\exists$ (18.4), it follows by the Rule of Substitution for Defined Formulas (160.3) that:

$$\Diamond\exists x\neg Fx \ \& \ \Diamond\exists xFx$$

Finally, by the commutativity of & (88.2.a), it follows that:

$$\Diamond\exists xFx \ \& \ \Diamond\exists x\neg Fx$$

($\leftarrow$) Reverse the reasoning. $\bowtie$

(202.3) ($\rightarrow$) Assume *Contingent*(F). Its definition (200.4) and the axiom $F\downarrow$ imply, by Rule $\equiv$S (91.1):

$$\neg(\mathit{Necessary}(F) \vee \mathit{Impossible}(F))$$

From this and the modally strict biconditional introduced by definition (200.1), we may infer, by the Rule of Substitution for Defined Formulas (160.3), that:

$$\neg(\Box\forall xFx \vee \mathit{Impossible}(F))$$

And by now familiar reasoning, definition (200.2) and the axiom $F\downarrow$ independently imply, by the special case of Rule $\equiv$S, the modally strict biconditional:

$$\mathit{Impossible}(F) \equiv \Box\forall x\neg Fx$$

So, by the Rule of Substitution (160.2), the last two displayed results imply:

$$(\vartheta) \quad \neg(\Box\forall xFx \vee \Box\forall x\neg Fx)$$

Now we know by (199.2) and (199.1), respectively, that the following are $\Box$-theorems: $\neg\overline{F}x \equiv Fx$ and $\overline{F}x \equiv \neg Fx$. Applying the commutativity of $\equiv$ to both, we therefore have, respectively, the following $\Box$-theorems:

$$\text{(a)} \quad Fx \equiv \neg\overline{F}x$$

$$\text{(b)} \quad \neg Fx \equiv \overline{F}x$$

Then (ϑ), (a) and (b) imply, by two applications of the Rule of Substitution (160.2) that:

(ξ) $\neg(\Box\forall x\neg\overline{F}x \vee \Box\forall x\overline{F}x)$

Now every instance of the commutativity of $\vee$ is a modally strict theorem. So from the relevant instance of the commutativity of $\vee$ and the Rule of Substitution (160.2) it follows from (ξ) that:

$$\neg(\Box\forall x\overline{F}x \vee \Box\forall x\neg\overline{F}x)$$

From this and the modally strict biconditional introduced by definition (200.1), we may infer, by the Rule of Substitution for Defined Formulas (160.3):

(ζ) $\neg(Necessary(\overline{F}) \vee \Box\forall x\neg\overline{F}x)$

Independently, definition (200.2) and the fact that $\overline{F}\downarrow$ imply, by the special case of Rule $\equiv$S:

$$Impossible(\overline{F}) \equiv \Box\forall x\neg\overline{F}x$$

So by the Rule of Substitution (160.2), (ζ) and the commuted form of this last result imply:

$$\neg(Necessary(\overline{F}) \vee Impossible(\overline{F}))$$

Hence, by conjoining the fact that $\overline{F}\downarrow$ with this last fact, it follows by definition (200.4) that $Contingent(\overline{F})$. ($\leftarrow$) Reverse the reasoning. ⋈

(203.1) By definition (200.1) and the Rule of Alphabetic Variants (114), it suffices to show $\Box\forall xLx$. From the fact that $[\lambda x\, E!x \rightarrow E!x]\downarrow$ (39.2), it follows from β-Conversion (48.2) that:

$$[\lambda x\, E!x \rightarrow E!x]x \equiv E!x \rightarrow E!x$$

Since the right side is a tautology, it follows that $[\lambda x\, E!x \rightarrow E!x]x$, i.e., Lx, by definition of L and Rule $=_{df}$I of Definiendum Introduction (120.2.b). Since this is a theorem, we may apply GEN and conclude $\forall xLx$. Since this is a $\Box$-theorem, it follows by RN that $\Box\forall xLx$. ⋈

(203.2) From (199.2) and the commutativity of $\equiv$, we know $Fx \equiv \neg\overline{F}x$. By two applications of GEN, it then follows that $\forall F\forall x(Fx \equiv \neg\overline{F}x)$. Since L clearly exists, given its definition, it follows that:

(a) $\forall x(Lx \equiv \neg\overline{L}x)$

Independently, from theorem (203.1), definition (200.1), and Rule $\equiv_{df}$E, it follows that $\Box\forall xLx$. So by the T schema (45.2):

(b) $\forall xLx$

By (99.3), it follows from (a) and (b) that:

$$\forall x \neg \overline{L}x$$

Since this is a $\Box$-theorem, it follows by RN that $\Box\forall x \neg \overline{L}x$. Since we also know $\overline{L}\downarrow$, it follows from definition (200.2) and Rule $\equiv_{df}$I that $Impossible(\overline{L})$. ⋈

(203.3) (Exercise)

(203.4) This follows either by theorems (203.3) and (202.1), or by theorem (203.2) and definition (200.3). ⋈

(203.5) (Exercise) [Use (203.3), (203.4), and facts about the distinctness of properties that are negations of one another, or use (195.1).]

(204.1) We may reason as follows:[446]

$$\begin{aligned} \Diamond\exists x(Fx \,\&\, \Diamond\neg Fx) &\equiv \exists x\Diamond(Fx \,\&\, \Diamond\neg Fx) && \text{by (167.3) and (167.4)} \\ &\equiv \exists x(\Diamond Fx \,\&\, \Diamond\neg Fx) && \text{(165.11) and (159.3)} \\ &\equiv \exists x(\Diamond\neg Fx \,\&\, \Diamond Fx) && \text{by (88.2.a) and (159.3)} \\ &\equiv \exists x\Diamond(\neg Fx \,\&\, \Diamond Fx) && \text{commute (165.11), and (159.3)} \\ &\equiv \Diamond\exists x(\neg Fx \,\&\, \Diamond Fx) && \text{by (167.4) and (167.3)} \end{aligned}$$ ⋈

(204.2) Apply (199.1), (199.2) and the Rule of Substitution (160.2) to the right-hand condition of (204.1). ⋈

(205.1) We want to show $\Diamond\exists x(E!x \,\&\, \Diamond\neg E!x)$. By CBF$\Diamond$ (167.4), it suffices to show $\exists x\Diamond(E!x \& \Diamond\neg E!x)$. By the commuted version of an appropriate instance of (165.11) and a Rule of Substitution (160.2), it suffices to show $\exists x(\Diamond E!x \& \Diamond\neg E!x)$. We prove this by first noting that by BF$\Diamond$ (167.3), axiom (45.4) implies $\exists x\Diamond(E!x \& \neg\mathcal{A}E!x)$. Suppose a is such an object, so that we know $\Diamond(E!a \,\&\, \neg\mathcal{A}E!a)$. By (162.3), this implies:

(ϑ) $\Diamond E!a \,\&\, \Diamond\neg\mathcal{A}E!a$

Now since $\neg\mathcal{A}E!a \equiv \mathcal{A}\neg E!a$ is an instance of a modally strict axiom (44.1), we may infer from (ϑ) by the Rule of Substitution (160.2) that:

(ξ) $\Diamond E!a \,\&\, \Diamond\mathcal{A}\neg E!a$

But as an instance of the commuted form of (164.4), we know $\Diamond\mathcal{A}\neg E!a \equiv \mathcal{A}\neg E!a$. From this last fact and (ξ) it follows by (88.4.f) that:

(ζ) $\Diamond E!a \,\&\, \mathcal{A}\neg E!a$

But the right conjunct of (ζ) implies $\Diamond\neg E!a$, by (164.3). So if we conjoin the left conjunct of (ζ) with this last fact, we have established $\Diamond E!a \,\&\, \Diamond\neg E!a$. Hence, $\exists x(\Diamond E!x \,\&\, \Diamond\neg E!x)$, which is what it sufficed to show. ⋈

[446] The following proof by Uri Nodelman simplified my original.

(205.2) Apply GEN to theorem (204.1), instantiate the result to $E!$ (which clearly exists), and then apply biconditional reasoning to the resulting instance and (205.1). ⋈

(205.3) By (205.1) and (168.4). ⋈

(205.4) By (205.2) and (168.4). ⋈

(205.5) By (202.2), (205.3) and (205.4). ⋈

(205.6) From (205.5) and (202.3). ⋈

(205.7) (Exercise)

(206.1) Assume $NonContingent(F)$. Then by definition (200.3):

(ϑ) $Necessary(F) \vee Impossible(F)$

Now assume, for reductio, that $\exists G(Contingent(G) \& G = F)$. Let P be an arbitrary such property, so that we have $Contingent(P) \& P = F$. Given the second conjunct and Rule =E, the first conjunct implies that $Contingent(F)$. Now, independently, from definition (200.4) and the axiom $F\downarrow$, it follows by Rule ≡S (91.1) that:

$$Contingent(F) \equiv \neg(Necessary(F) \vee Impossible(F))$$

Hence, $\neg(Necessary(F) \vee Impossible(F))$, which contradicts ($\vartheta$). ⋈

(206.2) (Exercise) [Hint: Use reasoning similar to that of (206.1).]

(206.3) (Exercise)

(207.1) ($\rightarrow$) Assume $NonContingent(p)$. Then by the 0-ary case of definition (200.3) and Rule $\equiv_{df}$E, we know $Necessary(p) \vee Impossible(p)$. From definition (200.1) and this last fact, it follows by (88.8.h) and biconditional reasoning that $\Box p \vee Impossible(p)$. Now, independently, definition (200.2) and the axiom $p\downarrow$ imply, by Rule ≡S (91.1), the biconditional $Impossible(p) \equiv \Box\neg p$. So by (88.8.g) and biconditional reasoning, it follows that:

(ϑ) $\Box p \vee \Box\neg p$

Independently, if we commute $\Box$-theorems (199.4) and (199.3), respectively, then we know that $p \equiv \neg\overline{p}$ and $\neg p \equiv \overline{p}$ are $\Box$-theorems. So by Rule RE (157.3), respectively:

(a) $\Box p \equiv \Box\neg\overline{p}$

(b) $\Box\neg p \equiv \Box\overline{p}$

From (a) and (ϑ), it follows by (88.8.h) that $\Box\neg\overline{p} \vee \Box\neg p$, and from (b) and this last result, it follows by (88.8.g) that $\Box\neg\overline{p} \vee \Box\overline{p}$. By the commutativity of $\vee$, this implies: $\Box\overline{p} \vee \Box\neg\overline{p}$. From definition (200.1) and this last fact, we may infer by (88.8.h) that $Necessary(\overline{p}) \vee \Box\neg\overline{p}$. Independently, by applying the special case of Rule $\equiv$S to definition (200.2) and the axiom $\overline{p}\downarrow$, we know the following is a modally strict theorem: $Impossible(\overline{p}) \equiv \Box\neg\overline{p}$. Hence, by now familiar reasoning, $Necessary(\overline{p}) \vee Impossible(\overline{p})$. So $NonContingent(\overline{p})$, by definition (200.3) and Rule $\equiv_{df}$I. ($\leftarrow$) Reverse the reasoning. $\bowtie$

(207.2) Assume $Contingent(p)$. Definition (200.4) and the axiom $p\downarrow$ then imply, by Rule $\equiv$S (91.1), $\neg(Necessary(p) \vee Impossible(p))$. From this and (the modally strict biconditional introduced by) definition (200.1), it follows by the Rule of Substitution for Defined Formulas (160.3) that $\neg(\Box p \vee Impossible(p))$. And since we know, by now familiar reasoning, that $Impossible(p) \equiv \Box\neg p$ is a modally strict theorem, it follows by the Rule of Substitution (160.2) that:

$$\neg(\Box p \vee \Box\neg p)$$

By De Morgan's Law (88.5.d), it follows that:

$$\neg\Box p \;\&\; \neg\Box\neg p$$

The modally strict theorem (158.11) and this last fact imply, by (88.4.e), that $\Diamond\neg p \;\&\; \neg\Box\neg p$. And from the definition of $\Diamond$ and this last fact, it follows by now familiar reasoning from (88.4.f) that:

$$\Diamond\neg p \;\&\; \Diamond p$$

Finally, by the commutativity of & (88.2.a), it follows that:

$$\Diamond p \;\&\; \Diamond\neg p$$

($\leftarrow$) Reverse the reasoning. $\bowtie$

(207.3) ($\rightarrow$) Assume $Contingent(p)$. Then its definition (200.4) and the axiom $p\downarrow$ imply, by the special case of Rule $\equiv$S, $\neg(Necessary(p) \vee Impossible(p))$. This and definition (200.1) imply, by the Rule of Substitution (160.2), $\neg(\Box p \vee Impossible(p))$. Since we know, by now familiar reasoning, that $Impossible(p) \equiv \Box\neg p$ is a modally strict theorem, it follows by the Rule of Substitution (160.2) that:

(ϑ) $\neg(\Box p \vee \Box\neg p)$

Independently, if we commute (199.4) and (199.3), we have as $\Box$-theorems:

(a) $p \equiv \neg\overline{p}$

(b) $\neg p \equiv \overline{p}$

From (ϑ) and (a), it follows by the Rule of Substitution (160.2) that $\neg(\Box\neg\overline{p} \vee \Box\neg p)$, and from this result and (b), it follows by the same rule that:

$$\neg(\Box\neg\overline{p} \vee \Box\overline{p})$$

So we may use an appropriate instance of the commutativity of $\vee$ (which is a $\Box$-theorem) and the Rule of Substitution (160.2) to transform the last formula into:

$$\neg(\Box\overline{p} \vee \Box\neg\overline{p})$$

From this and definition (200.1), we obtain $\neg(\mathit{Necessary}(\overline{p}) \vee \Box\neg\overline{p})$, by the Rule of Substitution (160.2). Independently, $\mathit{Impossible}(\overline{p}) \equiv \Box\neg\overline{p}$ is a modally strict theorem, by now familiar reasoning. By commuting this last result and combining it with the previous result, it follows by the Rule of Substitution (160.2) that:

$$\neg(\mathit{Necessary}(\overline{p}) \vee \mathit{Impossible}(\overline{p}))$$

Conjoining $\overline{p}\downarrow$ with this last result yields $\mathit{Contingent}(\overline{p})$, by (200.3). ($\leftarrow$) Reverse the reasoning. ⋈

(208.1) We're given the definition $p_0 =_{df} \forall x(E!x \rightarrow E!x)$. Now from $E!x \rightarrow E!x$, GEN and RN, we have $\Box\forall x(E!x \rightarrow E!x)$. By definition of p_0 and Rule $=_{df}$I, it follows that $\Box p_0$. Hence by definition (200.1) and Rule $\equiv_{df}$I, it follows that $\mathit{Necessary}(p_0)$. ⋈

(208.2) By the reasoning in (208.1), we established the following as a theorem:

(a) $\Box p_0$

Note that by the commutativity of $\equiv$, it follows from theorem (199.4) that $p \equiv \neg\overline{p}$, and hence by GEN that $\forall p(p \equiv \neg\overline{p})$. Since p_0 clearly exists, we can instantiate this last claim to p_0 to obtain $p_0 \equiv \neg\overline{p_0}$. From (a) and this last $\Box$-theorem, the Rule of Substitution (160.2) yields $\Box\neg\overline{p_0}$. We can conjoin the fact (exercise) that $\overline{p_0}\downarrow$ with this last result and apply definition (200.2) to conclude $\mathit{Impossible}(\overline{p_0})$. ⋈

(208.3) – **(208.5)** (Exercises)

(209)★ Suppose, for reductio, that $\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$. Let a be such an object, so that we know:

(ϑ) $E!a \,\&\, \neg\mathcal{A}E!a$

Then the second conjunct of (ϑ) implies $\mathcal{A}\neg E!a$, by the right-to-left direction of axiom (44.1). Then by axiom (43)★, it follows that $\neg E!a$, which contradicts the first conjunct of (ϑ).

(210.1) Suppose, for reductio, that $\mathcal{A}\exists x(E!x \mathbin{\&} \neg\mathcal{A}E!x)$. Then it follows by (139.10) that $\exists x\mathcal{A}(E!x \mathbin{\&} \neg\mathcal{A}E!x)$. Suppose a is such an object, so that we know $\mathcal{A}(E!a \mathbin{\&} \neg\mathcal{A}E!a)$. But this implies a contradiction, by the following reasoning:

$$\begin{array}{llll} \mathcal{A}(E!a \mathbin{\&} \neg\mathcal{A}E!a) & \rightarrow & \mathcal{A}E!a \mathbin{\&} \mathcal{A}\neg\mathcal{A}E!a & \text{by (139.2)} \\ & \rightarrow & \mathcal{A}E!a \mathbin{\&} \neg\mathcal{A}\mathcal{A}E!a & \text{commute (44.1), (88.4.f)} \\ & \rightarrow & \mathcal{A}E!a \mathbin{\&} \neg\mathcal{A}E!a & \text{commute (44.4), (160.2)} \qquad \bowtie \end{array}$$

(210.2) By (210.1), we know $\neg\mathcal{A}\exists x(E!x \mathbin{\&} \neg\mathcal{A}E!x)$. Hence, $\mathcal{A}\neg\exists x(E!x \mathbin{\&} \neg\mathcal{A}E!x)$, by (44.1). But then by (164.3), $\Diamond\neg\exists x(E!x \mathbin{\&} \neg\mathcal{A}E!x)$. ⋈

(210.3) Axiom (45.4) asserts $\Diamond\exists x(E!x \mathbin{\&} \neg\mathcal{A}E!x)$. By BF$\Diamond$ (167.3), it follows that $\exists x\Diamond(E!x \mathbin{\&} \neg\mathcal{A}E!x)$. Suppose a is such an object, so that we know $\Diamond(E!a \mathbin{\&} \neg\mathcal{A}E!a)$. Then by (162.3), it follows that both:

(ϑ) $\Diamond E!a$

(ξ) $\Diamond\neg\mathcal{A}E!a$

Now (ξ) implies $\neg\Box\mathcal{A}E!a$, by (158.11). But as an instance of the theorem (139.6), we know $\mathcal{A}E!a \equiv \Box\mathcal{A}E!a$. Hence, by biconditional syllogism, $\neg\mathcal{A}E!a$. Conjoining (ϑ) with this last result yields $\Diamond E!a \mathbin{\&} \neg\mathcal{A}E!a$. So $\exists x(\Diamond E!x \mathbin{\&} \neg\mathcal{A}E!x)$. ⋈

(211.1) – **(211.4)** (Exercises)

(212.1) Assume *NonContingent*(p). By the 0-ary case of definition (200.3), it follows that *Necessary*$(p) \vee$ *Impossible*(p). Now assume, for reductio, that $\exists q(\mathit{Contingent}(q) \mathbin{\&} q = p)$. Let q_1 be an arbitrary such proposition, so that *Contingent*$(q_1) \mathbin{\&} q_1 = p$. Then *Contingent*$(p)$. But, independently, from definition (200.4) and the fact that $p{\downarrow}$, it follows by Rule $\equiv$S that *Contingent*(p) is equivalent to $\neg(\mathit{Necessary}(p) \vee \mathit{Impossible}(p))$. Contradiction. ⋈.

(212.2) – **(212.4)** (Exercise)

(214.1) – **(214.2)** (Exercises)

(214.3) Assume *ContingentlyTrue*(p). Then by definition (213.1) and Rule $\equiv_{df}$E, we know:

(ϑ) $p \mathbin{\&} \Diamond\neg p$

By (213.2), we have to show: $\neg\overline{p} \mathbin{\&} \Diamond\overline{p}$. By (199.4), the first conjunct of (ϑ) implies $\neg\overline{p}$. So it remains to show $\Diamond\overline{p}$. Now if we commute (199.3), we obtain, as a modally strict theorem, $\neg p \equiv \overline{p}$. Hence it follows by the Rule of Substitution (160.2) from the second conjunct of (ϑ) that $\Diamond\overline{p}$. ⋈

(214.4) Assume *ContingentlyFalse*(p). Then by definition (213.2), we know:

(ϑ) $\neg p \mathbin{\&} \Diamond p$

By (213.1), we have to show: $\overline{p}$ & $\Diamond\neg\overline{p}$. By (199.3), the first conjunct of (ϑ) implies $\overline{p}$. So it remains to show $\Diamond\neg\overline{p}$. Now if we commute (199.4), we obtain a modally strict proof of $p \equiv \neg\overline{p}$. Hence it follows by the Rule of Substitution (160.2) from the second conjunct of (ϑ) that $\Diamond\neg\overline{p}$. ⋈

(214.5) Assume *ContingentlyTrue*(p) and *Necessary*(q). Assume for reductio that $p = q$. Then *Necessary*(p), i.e., $\Box p$, i.e., $\neg\Diamond\neg p$. But since p is contingently true, it follows by definition (213.1) that $\Diamond\neg p$. Contradiction. ⋈

(214.6) Assume *ContingentlyFalse*(p) and *Impossible*(q). Assume for reductio that $p = q$. Then *Impossible*(p), and so it follows from definition (200.4) *a fortiori* that $\Box\neg p$, i.e., $\neg\Diamond p$. But since p is contingently false, it follows by definition (213.2) that $\Diamond p$. Contradiction. ⋈

(215.1)★ By (209)★ and the definition of q_0 as $\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$, we know $\neg q_0$. By axiom (45.4) and the definition of q_0, we know $\Diamond q_0$. Hence, by &I and definition (213.2), it follows that *ContingentlyFalse*(q_0). ⋈

(215.2)★ (Exercise)

(217.1) Since we've defined $q_0 =_{df} \exists x(E!x \,\&\, \neg\mathcal{A}E!x)$, axiom (45.4) implies, by Rule $=_{df}$I:

(ϑ) $\Diamond q_0$

and theorem (210.2) implies:

(ξ) $\Diamond\neg q_0$

From these two facts, we can reason by cases from the tautology $q_0 \vee \neg q_0$ to the conclusion $\exists p$*ContingentlyTrue*(p).

Assume q_0. From this and (ξ), we have $q_0 \,\&\, \Diamond\neg q_0$. So, *ContingentlyTrue*(q_0), by definition (213.1). Hence $\exists p$*ContingentlyTrue*(p).

Assume $\neg q_0$. From this and (ϑ), we have $\neg q_0 \,\&\, \Diamond q_0$. So, *ContingentlyFalse*(q_0), by definition (213.2). But then, since $q_0\!\downarrow$, we may infer *ContingentlyTrue*($\overline{q_0}$), by (214.4). Since $\overline{q_0}\!\downarrow$, we may conclude $\exists p$*ContingentlyTrue*(p). ⋈

(217.2) If we let q_0 be $\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$, axiom (45.4) becomes:

(ϑ) $\Diamond q_0$

and theorem (210.2) becomes:

(ξ) $\Diamond\neg q_0$

From these two facts, we can reason by cases from the tautology $q_0 \vee \neg q_0$ to the conclusion $\exists p$*ContingentlyFalse*(p).

Assume q_0. From this and (ξ), we have $q_0 \,\&\, \Diamond\neg q_0$. So, *ContingentlyTrue*(q_0), by definition (213.1). But then, since $q_0\!\downarrow$, we may infer *ContingentlyFalse*($\overline{q_0}$), by (214.3). Since $\overline{q_0}\!\downarrow$, we may conclude $\exists p$*ContingentlyFalse*(p).

Assume $\neg q_0$. From this and (ϑ), we have $\neg q_0 \,\&\, \Diamond q_0$. So, *ContingentlyFalse*(q_0), by definition (213.2). Hence, $\exists p$*ContingentlyFalse*(p). ⋈

(217.2) [Simpler Proof] By (217.1), we know $\exists p$*ContingentlyTrue*(p). Let r be such a proposition, so that we know *ContingentlyTrue*(r). (See the discussion in Remark 218 as to why this assumption doesn't undermine the modal strictness of the reasoning.) Then since r exists by hypothesis, we may infer *ContingentlyFalse*($\overline{r}$), by (214.3). But $\overline{r}\!\downarrow$, and so, $\exists p$*ContingentlyFalse*(p). ⋈

(219.1) In the following, let:

$$Q_p =_{df} [\lambda z\, p]$$

Clearly, $Q_p\!\downarrow$, for every p. Before we begin our proof, note that the following, modally strict lemma governs Q_p:

(ϑ) $\forall p \forall x \Box (Q_p x \equiv p)$

> *Proof*. Since $[\lambda z\, p]\!\downarrow$, it follows by (183.2) that $\forall x([\lambda z\, p]x \equiv p)$ is a modally strict theorem. So by RN, $\Box\forall x([\lambda z p]x \equiv p)$ and by BF (167.1), $\forall x \Box([\lambda z p]x \equiv p)$. And by GEN, $\forall p \forall x \Box([\lambda z\, p]x \equiv p)$ is a theorem. Then by definition of Q_p, $\forall p \forall x \Box(Q_p x \equiv p)$.

Now to prove our theorem, by (217.1), we know that there are contingently true propositions. Let p_1 be such a proposition, so that we know $p_1 \,\&\, \Diamond\neg p_1$.[447] Now consider Q_{p_1}, which we know exists. We now argue as follows:

(A) Show: $p_1 \vdash Q_{p_1} y$. By (63.10), it suffices to show $p_1 \to Q_{p_1} y$. So assume p_1. Then by lemma (ϑ), $\forall x \Box (Q_{p_1} x \equiv p_1)$. Hence $\Box(Q_{p_1} y \equiv p_1)$, and by the T schema, $Q_{p_1} y \equiv p_1$. So $Q_{p_1} y$.

(B) Show: $\Diamond\neg p_1 \vdash \Diamond\neg Q_{p_1} y$. By (63.10), it suffices to show $\Diamond\neg p_1 \to \Diamond\neg Q_{p_1} y$. By instantiating p_1 and y into lemma (ϑ), we know $\Box(Q_{p_1} y \equiv p_1)$. *A fortiori*, $\Box(Q_{p_1} y \to p_1)$. This implies $\Box(\neg p_1 \to \neg Q_{p_1})$. Hence by K$\Diamond$, $\Diamond\neg p_1 \to \Diamond\neg Q_{p_1}$.

(C) Infer from (A) and (B):

[447] Again, we emphasize that although this assumption is not a necessary truth, it won't undermine the modal strictness of the reasoning, since the assumption will be discharged by $\exists$E and so the theorem will depend only on (217.1). See Remark (70) and the discussion in Remark (218). More specifically, we will derive the present theorem from the assumption that $p_1 \,\&\, \Diamond\neg p_1$. But it then follows by $\exists$E that our theorem is derivable from $\exists p(p \,\&\, \Diamond\neg p)$. Since latter is a modally strict theorem, by (217.1) and (213.1), our theorem will be proved from no assumptions. We develop the proof in detail below because it is important to see that it is modally strict, notwithstanding the introduction of assumptions that aren't necessary truths.

$$p_1 \,\&\, \Diamond\neg p_1 \vdash Qy \,\&\, \Diamond\neg Qy$$

by using the principle:

If $\varphi \vdash \psi$ and $\chi \vdash \theta$, then $\varphi \,\&\, \psi \vdash \chi \,\&\, \theta$

(D) By $\exists$I, we independently know: $Qy \,\&\, \Diamond\neg Qy \vdash \exists F\exists x(Fx \,\&\, \Diamond\neg Fx)$

(E) Hence from (C) and (D) it follows by (63.8) that:

$$p_1 \,\&\, \Diamond\neg p_1 \vdash \exists F\exists x(Fx \,\&\, \Diamond\neg Fx)$$

(F) It follows by $\exists$E that:

$$\exists p(p \,\&\, \Diamond\neg p) \vdash \exists F\exists x(Fx \,\&\, \Diamond\neg Fx)$$

(G) But by by (217.1) and (213.1), $\exists p(p \,\&\, \Diamond\neg p)$ is a theorem, and so it follows from (F) that $\exists F\exists x(Fx \,\&\, \Diamond\neg Fx)$, by (63.8).

Since the reasoning in (A) – (G) is straightforward, it remains only to show the principle used in (C):

If $\varphi \vdash \psi$ and $\chi \vdash \theta$, then $\varphi \,\&\, \psi \vdash \chi \,\&\, \theta$

Proof. Assume $\varphi \vdash \psi$ and $\chi \vdash \theta$. Then by the Deduction Theorem, it follows, respectively, that $\varphi \to \psi$ and $\chi \to \theta$. So by Double Composition (88.8.d), $(\varphi \,\&\, \chi) \to (\psi \,\&\, \theta)$. But then by (63.10), we have: $\varphi \,\&\, \chi \vdash \psi \,\&\, \theta$. ⋈

(219.2) (Exercise)

(221.1) In the following, let:

$$L =_{df} [\lambda x\, E!x \to E!x]$$

$$Q_p =_{df} [\lambda z\, p]$$

Clearly, L exists, and Q_p exists for every proposition p. Before starting the proof proper, we first prove two modally strict lemmas:

(ϑ) $\forall p(p \to \forall x(Lx \equiv Q_p x))$

Proof. By GEN, we show $p \to \forall x(Lx \equiv Q_p x)$. So assume p. By GEN, it then suffices to show $Lx \equiv Q_p x$. Since $L\!\downarrow$ and $Q_p\!\downarrow$, we independently know, by Strengthened β-Conversion (181) and the definitions of L and Q_p:

(ζ) $Lx \equiv (E!x \to E!x)$

(ω) $Q_p x \equiv p$

Hence, to show $Lx \equiv Q_p x$, it suffices to show $(E!x \rightarrow E!x) \equiv p$. But the left side is a tautology (74) and the right side is true by assumption. So since both sides of the biconditional are true, the biconditional is true.

(ξ) $\forall p(\Diamond\neg p \rightarrow \Diamond\neg\forall x(Lx \equiv Q_p x))$

Proof. By GEN, we show $\Diamond\neg p \rightarrow \Diamond\neg\forall x(Lx \equiv Q_p x)$. But we can infer this by RM$\Diamond$ (157.2) from a modally strict proof of $\neg p \rightarrow \neg\forall x(Lx \equiv Q_p x)$. So assume $\neg p$. For reductio, assume $\forall x(Lx \equiv Q_p x)$. Then, $Lx \equiv Q_p x$. From (ζ) and (ω) above (which were established independently) and this last result it follows that:

$$(E!x \rightarrow E!x) \equiv p$$

But since $\neg p$ by assumption, it follows that $\neg(E!x \rightarrow E!x)$, which contradicts the tautology $E!x \rightarrow E!x$ (74).

Now to prove our theorem. By (217.1), we know that there are contingently true propositions. Let p_1 be such a proposition, so that we know $p_1 \,\&\, \Diamond\neg p_1$.[448] We now show that L and Q_{p_1} are the witnesses to the theorem we're trying to prove. The proof then, goes as follows:

(A) Show: $p_1 \vdash \forall x(Lx \equiv Q_{p_1} x)$. By (63.10), it suffices to show $p_1 \rightarrow \forall x(Lx \equiv Q_{p_1} x)$. This, however, is an instance of lemma (ϑ).

(B) Show: $\Diamond\neg p_1 \vdash \Diamond\neg\forall x(Lx \equiv Q_{p_1} x)$. But it suffices by (63.10) to show $\Diamond\neg p_1 \rightarrow \Diamond\neg\forall x(Lx \equiv Q_{p_1} x)$. This, however, is an instance of lemma (ξ).

(C) It follows from (A) and (B) that:[449]

$$p_1 \,\&\, \Diamond\neg p_1 \vdash \forall x(Lx \equiv Q_{p_1} x) \,\&\, \Diamond\neg\forall x(Lx \equiv Q_{p_1} x)$$

(D) But clearly, by $\exists$I, we know the following about the conclusion of the derivation in (C):

$$\forall x(Lx \equiv Q_{p_1} x) \,\&\, \Diamond\neg\forall x(Lx \equiv Q_{p_1} x) \vdash \exists F\exists G(\forall x(Fx \equiv Gx) \,\&\, \Diamond\neg\forall x(Fx \equiv Gx))$$

(E) Hence, from (C) and (D) it follows by (63.8) that:

$$p_1 \,\&\, \Diamond\neg p_1 \vdash \exists F\exists G(\forall x(Fx \equiv Gx) \,\&\, \Diamond\neg\forall x(Fx \equiv Gx))$$

[448] See footnote 447.

[449] The principle used to go from (A) and (B) to (C) is:

If $\varphi \vdash \psi$ and $\chi \vdash \theta$, then $\varphi \,\&\, \psi \vdash \chi \,\&\, \theta$

To see that this holds, assume $\varphi \vdash \psi$ and $\chi \vdash \theta$. Then by the Deduction Theorem, it follows, respectively, that $\varphi \rightarrow \psi$ and $\chi \rightarrow \theta$. So by Double Composition (88.8.d), $(\varphi \,\&\, \chi) \rightarrow (\psi \,\&\, \theta)$. But then by (63.10), we have: $\varphi \,\&\, \chi \vdash \psi \,\&\, \theta$.

(F) It follows by ∃E that:

$$\exists p(p \,\&\, \Diamond\neg p) \vdash \exists F \exists G(\forall x(Fx \equiv Gx) \,\&\, \Diamond\neg\forall x(Fx \equiv Gx))$$

(G) But by (217.1) and (213.1), $\exists p(p \,\&\, \Diamond\neg p)$ is a theorem, and so it follows from (F) that $\vdash \exists F \exists G(\forall x(Fx \equiv Gx) \,\&\, \Diamond\neg\forall x(Fx \equiv Gx))$, by (63.8). ⋈

(221.2) In the following, let:

$L =_{df} [\lambda x\, E!x \to E!x]$

$Q_p =_{df} [\lambda z\, p]$

Clearly, L exists, and Q_p exists for every proposition p. Before starting the proof proper, we first prove two modally strict lemmas:

(ϑ) $\forall p(\neg p \to \neg\forall x(Lx \equiv Q_p x))$

Proof. By GEN, we show: $\neg p \to \neg\forall x(Lx \equiv Q_p x)$. Assume $\neg p$. For reductio, assume $\forall x(Lx \equiv Q_p x)$. But by the definition of L, we know $\forall x Lx$. So it follows that $\forall x Q_p x$. Hence, by definition of Q_p, $\forall x([\lambda z\, p]x)$. But then $[\lambda x\, p]x$, by ∀E, and by β-Conversion, p. Contradiction.

(ξ) $\forall p(\Diamond p \to \Diamond\forall x(Lx \equiv Q_p x))$

Proof. By GEN, we show: $\Diamond p \to \Diamond\forall x(Lx \equiv Q_p x)$. But we can infer this by RM◊ (157.2) from a modally strict proof of $p \to \forall x(Lx \equiv Q_p x)$. So assume p. Then, by β-Conversion, $[\lambda z\, p]x$. Hence, by definition of Q_p, $Q_p x$. Since x doesn't occur free in any assumption, $\forall x Q_p x$, by GEN. But we know $\forall x Lx$, by definition and properties of L. Hence, $\forall x(Lx \equiv Q_p x)$.

Now we prove our theorem. By (217.2), we know that there are contingently false propositions. Let p_1 be such a proposition, so that we know $\neg p_1 \,\&\, \Diamond p_1$. Now we are going to show that L and Q_{p_1} are the witnesses to the theorem we're trying to prove. The proof, then, goes as follows:

(A) Show: $\neg p_1 \vdash \neg\forall x(Lx \equiv Q_{p_1} x)$. By (63.10), it suffices to show $\neg p_1 \to \neg\forall x(Lx \equiv Q_{p_1} x)$. This, however, is an instance of lemma (ϑ).

(B) Show: $\Diamond p_1 \vdash \Diamond\forall x(Lx \equiv Q_{p_1} x)$. But it suffices, by (63.10), to show $\Diamond p_1 \vdash \Diamond\forall x(Lx \equiv Q_{p_1} x)$. This, however, is an instance of lemma (ξ).

(C) Then it follows from (A) and (B), by the reasoning in footnote 449, that:

$$\neg p_1 \,\&\, \Diamond p_1 \vdash \neg\forall x(Lx \equiv Q_{p_1} x) \,\&\, \Diamond\forall x(Lx \equiv Q_{p_1} x)$$

(D) But clearly, by ∃I, we also know the following about the conclusion of the derivation in (C):

$$\neg\forall x(Lx \equiv Q_{p_1}x) \,\&\, \Diamond\forall x(Lx \equiv Q_{p_1}x) \vdash \exists F\exists G(\neg\forall x(Fx \equiv Gx) \,\&\, \Diamond\forall x(Fx \equiv Gx))$$

(E) Hence, from (C) and (D) it follows by (63.8) that:

$$\neg p_1 \,\&\, \Diamond p_1 \vdash \exists F\exists G(\neg\forall x(Fx \equiv Gx) \,\&\, \Diamond\forall x(Fx \equiv Gx))$$

(F) It follows by $\exists$E that:

$$\exists p(\neg p \,\&\, \Diamond p) \vdash \exists F\exists G(\neg\forall x(Fx \equiv Gx) \,\&\, \Diamond\forall x(Fx \equiv Gx))$$

(G) But by (217.1) and (213.1), $\exists p(\neg p \,\&\, \Diamond p)$ is a theorem, and so it follows from (F) that $\vdash \exists F\exists G(\neg\forall x(Fx \equiv Gx) \,\&\, \Diamond\forall x(Fx \equiv Gx))$, by (63.8). ⋈

(221.3) In the following, let:

$$L =_{df} [\lambda x\, E!x \rightarrow E!x]$$

$$Q_p =_{df} [\lambda z\, p]$$

Clearly, L exists, and Q_p exists for every proposition p. Before starting the proof proper, we first prove two modally strict lemmas:

(ϑ) $\forall p(\mathcal{A}\neg p \rightarrow \mathcal{A}\neg\forall x(Lx \equiv Q_p x))$

Proof. In the proof of (221.2), we independently established, as lemma (ϑ), that $\forall p(\neg p \rightarrow \neg\forall x(Lx \equiv Q_p x))$. Since this is a theorem, it follows by the Rule of Actualization that $\mathcal{A}\forall p(\neg p \rightarrow \neg\forall x(Lx \equiv Q_p x))$. Hence by axiom (44.3), $\forall p\mathcal{A}(\neg p \rightarrow \neg\forall x(Lx \equiv Q_p x))$. But as an instance of axiom (44.2), the following is a modally strict theorem:

$$\mathcal{A}(\neg p \rightarrow \neg\forall x(Lx \equiv Q_p x)) \equiv (\mathcal{A}\neg p \rightarrow \mathcal{A}\neg\forall x(Lx \equiv Q_p x))$$

So (ϑ) follows by a Rule of Substitution.

(ξ) $\forall p(\Diamond p \rightarrow \Diamond\forall x(Lx \equiv Q_p x))$

Proof. See the proof of (ξ) in (221.2).

Now we prove our theorem. Let q_0 again be the proposition $\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$. Then it follows from (210.1) by axiom (44.1) that $\mathcal{A}\neg q_0$. Independently, by axiom (45.4), we know $\Diamond q_0$. So both $\mathcal{A}\neg q_0$ and $\Diamond q_0$ are theorems.

Now consider Q_{q_0}. We now show that L and Q_{q_0}, are the witnesses to the theorem we're trying to prove. The proof, then, goes as follows:

(A) Show: $\mathcal{A}\neg q_0 \vdash \mathcal{A}\neg\forall x(Lx \equiv Q_{q_0}x)$. By (63.10), it suffices to show $\mathcal{A}\neg q_0 \rightarrow \mathcal{A}\neg\forall x(Lx \equiv Q_{q_0}x)$. This, however, is an instance of lemma (ϑ).

(B) Show: $\Diamond q_0 \vdash \Diamond\forall x(Lx \equiv Q_{q_0}x)$. But it suffices, by (63.10), to show $\Diamond q_0 \vdash \Diamond\forall x(Lx \equiv Q_{q_0}x)$. This, however, is an instance of lemma (ξ).

(C) Then it follows from (A) and (B), by reasoning analogous to (but not exactly the same as) that in footnote 449, that:

$$\mathcal{A}\neg q_0, \Diamond q_0 \vdash \mathcal{A}\neg\forall x(Lx \equiv Q_{q_0}x) \,\&\, \Diamond\forall x(Lx \equiv Q_{q_0}x)$$

(D) But clearly, by ∃I, we also know the following about the conclusion of the derivation in (C):

$$\mathcal{A}\neg\forall x(Lx \equiv Q_{q_0}x) \,\&\, \Diamond\forall x(Lx \equiv Q_{q_0}x) \vdash \exists F\exists G(\mathcal{A}\neg\forall x(Fx \equiv Gx) \,\&\, \Diamond\forall x(Fx \equiv Gx))$$

(E) Hence, from (C) and (D) it follows by (63.8) that:

$$\mathcal{A}\neg q_0, \Diamond q_0 \vdash \exists F\exists G(\mathcal{A}\neg\forall x(Fx \equiv Gx) \,\&\, \Diamond\forall x(Fx \equiv Gx))$$

(F) Since we established (at the outset) that both $\mathcal{A}\neg q_0$ and $\Diamond q_0$ are theorems, it follows that $\vdash \exists F\exists G(\mathcal{A}\neg\forall x(Fx \equiv Gx) \,\&\, \Diamond\forall x(Fx \equiv Gx))$, by a generalization of (63.8) when Γ is empty.[450] ⋈

(221.4) By GEN, it suffices to show:

(ϑ) $\exists G(\forall x(Fx \equiv Gx) \,\&\, \Diamond\neg\forall x(Fx \equiv Gx))$

We first prove some facts before laying out the proof of (ϑ).[451]

Let q be any proposition you please and consider the following two properties, both of which exist, by (39.2):

$$F_q =_{df} [\lambda z\, Fz \,\&\, q]$$

$$F'_q =_{df} [\lambda z\, (Fz \,\&\, q) \vee \neg q]$$

Given these definitions, it follows that:

(A) $q \rightarrow \forall x(Fx \equiv F_q x)$

(B) $q \rightarrow \forall x(Fx \equiv F'_q x)$

(C) $\Diamond\neg q \rightarrow \Diamond\neg\forall x(F_q x \equiv F'_q x)$

(D) $\Box\forall x(Fx \equiv F_q x) \rightarrow (\Diamond\neg\forall x(F_q x \equiv F'_q x) \equiv \Diamond\neg\forall x(Fx \equiv F'_q x))$

We leave (A) and (B) as exercises. To see that (C) holds, note that without too much work, one can show:

$$\neg q \rightarrow \neg\forall x(F_q x \equiv F'_q x)$$

[450] The generalization in question, which we leave as an exercise, states: if $\vdash \varphi_1$ and ... and $\vdash \varphi_n$ and $\varphi_1, \ldots, \varphi_n \vdash \psi$, then $\vdash \psi$.

[451] I'm indebted to Uri Nodelman for suggesting this proof strategy.

> *Proof*. Assume $\neg q$ and, for reductio, $\forall x(F_q x \equiv F'_q x)$, so that we know $F_q x \equiv F'_q x$. From this last fact, it follows by definitions of F_q, F'_q and β-Conversion that $(Fx \,\&\, q) \equiv ((Fx \,\&\, q) \vee \neg q)$. But our assumption $\neg q$ implies $(Fx \,\&\, q) \vee \neg q$. And from this and the biconditional just established, it follows that (Fx and) q. Contradiction.

Since the fact we just established is modally strict, it follows by RM$\Diamond$ that (C).

To see that (D) holds, assume $\Box\forall x(Fx \equiv F_q x)$, which by the commutativity of the biconditional and easy quantified modal logic reasoning implies $\Box\forall x(F_q x \equiv Fx)$. Note independently by reasoning from the tautology (88.4.a), it is easy to establish, as a modally strict theorem, that:

$$\forall\alpha(\varphi \equiv \psi) \to (\forall\alpha(\varphi \equiv \chi) \equiv \forall\alpha(\psi \equiv \chi))$$

But the consequent of this conditional, i.e., $\forall\alpha(\varphi \equiv \chi) \equiv \forall\alpha(\psi \equiv \chi)$ implies, by modally strict means, $\neg\forall\alpha(\varphi \equiv \chi) \equiv \neg\forall\alpha(\psi \equiv \chi)$. So by hypothetical syllogism, the following is a modally strict theorem:

$$\forall\alpha(\varphi \equiv \psi) \to (\neg\forall\alpha(\varphi \equiv \chi) \equiv \neg\forall\alpha(\psi \equiv \chi))$$

From this it follows by RM that:

$$\Box\forall\alpha(\varphi \equiv \psi) \to \Box(\neg\forall\alpha(\varphi \equiv \chi) \equiv \neg\forall\alpha(\psi \equiv \chi))$$

As a particular instance of this, we have:

$$\Box\forall x(F_q x \equiv Fx) \to \Box(\neg\forall x(F_q x \equiv F'_q x) \equiv \neg\forall x(Fx \equiv F'_q x))$$

Since we've already established the antecedent, it follows that:

$$(\zeta)\quad \Box(\neg\forall x(F_q x \equiv F'_q x) \equiv \neg\forall x(Fx \equiv F'_q x))$$

Now it is straightforward to derive from theorem (158.4) that:

$$\Box(\varphi \equiv \psi) \to (\Diamond\varphi \equiv \Diamond\psi)$$

By the relevant instance of this last fact, it follows from (ζ) that:

$$\Diamond\neg\forall x(F_q x \equiv F'_q x) \equiv \Diamond\neg\forall x(Fx \equiv F'_q x)$$

concluding our proof of (D). And with this proof of (D), we may conclude our preliminaries by noting that one may universally generalize on the free variable q in (A) – (D), by GEN.

Now to prove (ϑ). We know that there is a contingently true proposition (217.1). Then let p_1 be such a proposition, so that p_1 and $\Diamond\neg p_1$. Now consider F_{p_1} and F'_{p_1}. Either $\Box\forall x(Fx \equiv F_{p_1}x)$ or not. We show that there is a witness to (ϑ) in both cases.

Case 1. $\Box\forall x(Fx \equiv F_{p_1}x)$. Note that it follows from the fact that p_1, by (B), that $\forall x(Fx \equiv F'_{p_1}x)$. So to show that F'_{p_1} is a witness to (ϑ), it remains only to show $\Diamond\neg\forall x(Fx \equiv F'_{p_1}x)$. Now from the fact that $\Diamond\neg p_1$, it follows from (C) that:

(ξ) $\Diamond\neg\forall x(F_{p_1}x \equiv F'_{p_1}x)$

But since we're in Case 1, $\Box\forall x(Fx \equiv F_{p_1}x)$. This implies, by instantiating (D) to p_1, that:

$(\zeta)'$ $\Diamond\neg\forall x(F_{p_1}x \equiv F'_{p_1}x) \equiv \Diamond\neg\forall x(Fx \equiv F'_{p_1}x)$

But (ξ) and $(\zeta)'$ imply $\Diamond\neg\forall x(Fx \equiv F'_{p_1}x)$.

Case 2. $\Box\forall x(Fx \equiv F_{p_1}x)$. Then, $\Diamond\neg\forall x(Fx \equiv F_{p_1}x)$. Note that it follows from the fact that p_1, by (A), that $\forall x(Fx \equiv F_{p_1}x)$. So F_{p_1} is a witness to (ϑ).[452] ⋈

(221.5) – **(221.6)** (Exercises)

(222.1) By definition (24), we have to show $\neg(O! = A!)$. For reductio, assume $O! = A!$. Now, independently, by (115.3), we know $O!x \equiv \Diamond E!x$. Hence by Rule =E, $A!x \equiv \Diamond E!x$. But we also independently know, by (115.4), that $A!x \equiv \neg\Diamond E!x$. Hence, $\Diamond E!x \equiv \neg\Diamond E!x$. Contradiction. ⋈

(222.2) By (115.3), we know $O!x \equiv \Diamond E!x$. It follows from this, by the relevant instance of (88.3.b) and biconditional syllogism that $O!x \equiv \neg\neg\Diamond E!x$. But since the modally strict theorem (115.4) implies, by the commutativity of the biconditional, that $\neg\Diamond E!x \equiv A!x$ is a modally strict theorem, it follows by a Rule of Substitution (160.2) that $O!x \equiv \neg A!x$. ⋈

(222.3) (Exercise)

(222.4) By (202.2) and &I, it suffices to establish $\Diamond\exists xO!x$ and $\Diamond\exists x\neg O!x$.

To prove $\Diamond\exists xO!x$, we start with theorem (205.3), i.e., $\Diamond\exists xE!x$. By BF$\Diamond$ (167.3), we obtain $\exists x\Diamond E!x$. Note independently that $[\lambda x\,\Diamond E!x]$ exists. So it follows from β-Conversion and the commutativity of $\equiv$ that $\Diamond E!x \equiv [\lambda x\,\Diamond E!x]x$. Since this is a $\Box$-theorem, it follows from $\exists x\Diamond E!x$ that $\exists x([\lambda x\,\Diamond E!x]x)$, by the Rule of Substitution (160.2). But by definition of $O!$ (22.1), this is just $\exists xO!x$. So by T$\Diamond$ (163.1), we have $\Diamond\exists xO!x$.

To prove $\Diamond\exists x\neg O!x$, let φ be an formula in which x doesn't occur free. Then by comprehension axiom for abstract objects (53), we know $\exists x(A!x \,\&\, \forall F(xF \equiv \varphi))$. By (103.5), it follows that $\exists xA!x$. Now we know by the previous theorem (222.3) that $A!x \equiv \neg O!x$. This is a $\Box$-theorem and so by the Rule of Substitution (160.2), we may infer $\exists x\neg O!x$ from $\exists xA!x$. But by T$\Diamond$, it then follows that $\Diamond\exists x\neg O!x$. ⋈

(222.5) – **(222.7)** (Exercises)

(222.8) From (222.4) and (202.3). ⋈

[452] I'm indebted to Daniel Kirchner and his implementation in Isabelle/HOL (2017 [2021] and 2022) for finding an invalid application of a Rule of Necessary Equivalence in an earlier version of this proof. The rule was mistakenly applied under an assumption in the proof of fact (D).

(222.9) From (222.5) and (202.3). ⋈

(224.1) (→) Assume *WeaklyContingent*(F). Then by definition (223) and Rule $\equiv_{df}$E, we know both *Contingent*(F) and $\forall x(\Diamond Fx \rightarrow \Box Fx)$. From the former, it follows that *Contingent*$(\overline{F})$, by (202.3). So, to show *WeaklyContingent*$(\overline{F})$, it remains, by Rule $\equiv_{df}$I, to establish $\forall x(\Diamond \overline{F}x \rightarrow \Box \overline{F}x)$. By GEN, it suffices to show $\Diamond \overline{F}x \rightarrow \Box \overline{F}x$. So assume $\Diamond \overline{F}x$. Since it is a $\Box$-theorem (199.1) that $\overline{F}x \equiv \neg Fx$, it follows by the Rule of Substitution (160.2) that $\Diamond \neg Fx$, i.e., $\neg \Box Fx$. Now we want to show $\Box \overline{F}x$, but for reductio, assume $\neg \Box \overline{F}x$, i.e., by definition of relation negation and Rule $=_{df}$E, $\neg \Box [\lambda y\, \neg Fy]x$. Since $[\lambda y\, \neg Fy]\downarrow$, Strengthened β-Conversion implies the $\Box$-theorem $[\lambda y\, \neg Fy]x \equiv \neg Fx$. So it follows by the Rule of Substitution (160.2) that $\neg \Box \neg Fx$, i.e., $\Diamond Fx$, by definition of $\Diamond$. Since we already know $\forall x(\Diamond Fx \rightarrow \Box Fx)$, it follows that $\Box Fx$. Contradiction. (←) Exercise. ⋈

(224.2) (Exercise)

(225.1) By (222.4), we know *Contingent*$(O!)$. By the left-to-right direction of (180.5) and GEN, we know $\forall x(\Diamond O!x \rightarrow \Box O!x)$. Hence, by definition (223) and Rule $\equiv_{df}$I, *WeaklyContingent*$(O!)$. ⋈

(225.2) By (222.5), we know *Contingent*$(A!)$. By the left-to-right direction of (180.6) and GEN, we know: $\forall x(\Diamond A!x \rightarrow \Box A!x)$. Hence by definition (223) and Rule$\equiv_{df}$I, *WeaklyContingent*$(A!)$. ⋈

(225.3) By (223), we have to show either $\neg$*Contingent*$(E!)$ or $\neg\forall x(\Diamond E!x \rightarrow \Box E!x)$. But we already know, by (205.5), that *Contingent*$(E!)$. So we have to show $\neg\forall x(\Diamond E!x \rightarrow \Box E!x)$. For reductio, assume $\forall x(\Diamond E!x \rightarrow \Box E!x)$. Now axiom (45.4) tells us $\Diamond \exists x(E!x \,\&\, \neg \mathcal{A}E!x)$. So by BF$\Diamond$ (167.3), $\exists x \Diamond(E!x \,\&\, \neg \mathcal{A}E!x)$. Suppose a is such an object, so that we know $\Diamond(E!a \,\&\, \neg \mathcal{A}E!a)$. Then by (162.3):

(ϑ) $\Diamond E!a \,\&\, \Diamond \neg \mathcal{A}E!a)$

The first conjunct of (ϑ) implies, by our reductio assumption, that $\Box E!a$. But from this, it follows by (132) that $\mathcal{A}E!a$, and by axiom (46.1), that $\Box \mathcal{A}E!a$. But the second conjunct of (ϑ) implies that $\neg \Box \mathcal{A}E!a$ (158.11). Contradiction. ⋈

(225.4) (Exercise) [Hint: Show that L fails the first conjunct in the definition of *weakly contingent*. Appeal to previous theorems.] ⋈

(225.5) (Exercise)

(225.6) (Exercise)

(226) Consider the following entities and definitions:[453]

[453] The following extended proof sketch was contributed by Daniel Kirchner, personal communication, 29 May 2018, and reproduced with permission.

- By (210.3), we know $\exists x(\Diamond E!x \,\&\, \neg \mathcal{A}E!x)$. So let o be an (arbitrary) such object, so that we know $\Diamond E!o \,\&\, \neg \mathcal{A}E!o$.
- Let a be an (arbitrary) abstract object.
- Let $\triangle$ be the *necessary or not-actually-true-but-possible* operator,[454] defined as follows:

$$\triangle\varphi \equiv_{df} \Box\varphi \vee (\neg\mathcal{A}\varphi \,\&\, \Diamond\varphi)$$

- Define $q_0 =_{df} \exists x(E!x \,\&\, \neg\mathcal{A}E!x)$, so that we know $\Diamond q_0$ by axiom (45.5), $\neg\mathcal{A}q_0$ by theorem (210.1), and $\neg\Box q_0$ by (210.2) and modal negation.
- Define $L =_{df} [\lambda x\, E!x \to E!x]$.

Now consider Figure 1 (see below). In this figure, each cell represents a for-

	$\mathcal{A}\Pi o$	$\triangle\Pi o$	$\mathcal{A}\Pi a$	$\triangle\Pi a$	entries as binary
$\overline{L}$	–	–	–	–	0
$[\lambda y\, A!y \,\&\, q_0]$	–	–	–	+	1
$[\lambda y\, A!y \,\&\, \overline{q_0}]$	–	–	+	–	2
$A!$	–	–	+	+	3
$E!$	–	+	–	–	4
$[\lambda y\, q_0]$	–	+	–	+	5
$[\lambda y\, E!y \vee (A!y \,\&\, \overline{q_0})]$	–	+	+	–	6
$[\lambda y\, A!y \vee E!y]$	–	+	+	+	7
$[\lambda y\, O!y \,\&\, \overline{E!}y]$	+	–	–	–	8
$[\lambda y\, \overline{E!}y \,\&\, (O!y \vee q_0)]$	+	–	–	+	9
$[\lambda y\, \overline{q_0}]$	+	–	+	–	10
$\overline{E!}$	+	–	+	+	11
$O!$	+	+	–	–	12
$[\lambda y\, O!y \vee q_0]$	+	+	–	+	13
$[\lambda y\, O!y \vee \overline{q_0}]$	+	+	+	–	14
L	+	+	+	+	15

Figure 1: Assign + if Provable and – if Negation is Provable

mula, namely, the result of taking the expression labeling a cell's row as the value of the metavariable Π labeling a cell's column. We place '+' in a cell

[454] Since we know by (217.1) and (217.2) that there are contingently true and contingently false propositions, then we intuitively know (in terms of semantically primitive possible worlds) that there is a non-actual possible world distinct from the actual world. (This will be proved *within* our system in Chapter 12.) So, given the following definition, $\triangle\varphi$ holds if and only if either φ is true at every possible world or φ true in some non-actual world but not at the actual world.

to indicate that the formula represented is provable, and '−' to indicate that the negation of the formula is provable. To see that the rows are pairwise distinct, note that the entries for a given row can be interpreted as the 4-bit binary representation (reading − as 0 and + as 1) of a number between 0 to 15.

Note the following two facts:

- as a general truth about object theory, any two formulas φ and ψ are equivalent if one results from the other by the substitution of identicals, and
- since no two rows in the table are the same, all of the properties are pairwise distinct; if two properties F and G in the left column were such that $F = G$, then substituting G for F would result in an equivalent formula and so the rows for F and G in the table would have to be the same.

Now, by straightforward arguments, one can prove that each cell in the table is correctly labeled with + or −, if given the following facts about particular properties and propositions (and the dual facts about their negations):

- $\Diamond q_0$, $\neg\mathcal{A}q_0$, $\neg\Box q_0$, $\Diamond\neg\overline{q_0}$, $\mathcal{A}\overline{q_0}$, $\Diamond\overline{q_0}$
- $\Box\forall x\, Lx$, $\Box\forall x\neg\overline{L}x$
- $O!o$, $\Box O!o$, $\Box\neg A!o$, $\triangle E!o$
- $\Box A!a$, $\Box\neg O!a$, $\Box\neg E!a$
- $\forall x(\triangle[\lambda y\, \overline{q_0}]x)$
- $\forall x(\neg\mathcal{A}[\lambda y\, \overline{q_0}]x)$

and the general principles:

- $\Box\varphi \to \mathcal{A}\varphi$ (132)
- $\Box\varphi \to \triangle\varphi$

It is then straightforward to show that the cells in the rows of $A!$, $E!$, $O!$, L and $[\lambda y\, \overline{q_0}]$ are correctly labeled, i.e., to produce the relevant proofs. The label for each cell in the remaining rows is just a consequence of the labels for the cells in the first five rows (by applying combinatory methods).

It may make it easier to see that the cells are correctly labeled by regrouping the properties, so that it becomes clearer that the cells in the final 11 rows are negations and combinations of (negations of) the cells in the first 5 rows. We've done this in Figure 2 (below), though the regrouping in this figure may make it more difficult to see that all of the properties are pairwise distinct, since the rows are no longer systematically ordered by their entries. ⋈

	$\mathscr{A}\Pi o$	$\triangle\Pi o$	$\mathscr{A}\Pi a$	$\triangle\Pi a$	entries as binary
L	+	+	+	+	15
$A!$	−	−	+	+	3
$O!$	+	+	−	−	12
$E!$	−	+	−	−	4
$[\lambda y\, \overline{q_0}]$	+	−	+	−	10
$\overline{L}$	−	−	−	−	0
$\overline{E!}$	+	−	+	+	11
$[\lambda y\, q_0]$	−	+	−	+	5
$[\lambda y\, A!y \vee E!y]$	−	+	+	+	7
$[\lambda y\, O!y \,\&\, \overline{E!}y]$	+	−	−	−	8
$[\lambda y\, O!y \vee \overline{q_0}]$	+	+	+	−	14
$[\lambda y\, O!y \vee q_0]$	+	+	−	+	13
$[\lambda y\, A!y \,\&\, q_0]$	−	−	−	+	1
$[\lambda y\, A!y \,\&\, \overline{q_0}]$	−	−	+	−	2
$[\lambda y\, \overline{E!}y \,\&\, (O!y \vee q_0)]$	+	−	−	+	9
$[\lambda y\, E!y \vee (A!y \,\&\, \overline{q_0})]$	−	+	+	−	6

Figure 2: Proof of Pairwise Distinctness

(227.1) In the first part of the proof of (222.4), we proved $\exists xO!x$ along the way, by modally strict means. So our theorem follows by RN. ⋈

(227.2) In the second part of the proof of (222.4), we proved $\exists xA!x$ along the way, by modally strict means. So our theorem follows by RN. ⋈

(227.3) From the modally strict theorem (222.3), i.e., $A!x \equiv \neg O!x$, and the previous theorem (227.2), it follows, by the Rule of Substitution (160.2), that $\Box\exists x\neg O!x$. Hence from this and the modally strict theorem (103.2), i.e., $\exists x\neg\varphi \equiv \neg\forall x\varphi$, it follows that $\Box\neg\forall xO!x$, also by the Rule of Substitution (160.2). ⋈

(227.4) From modally strict theorem (222.2), i.e., that $O!x \equiv \neg A!x$, and theorem (227.1), it follows, by the Rule of Substitution (160.2), that $\Box\exists x\neg A!x$. From this and the modally strict theorem (103.2), it follows by the Rule of Substitution (160.2) that $\Box\neg\forall xA!x$. ⋈

(227.5) In the second part of the proof of (222.4), we established $\exists xA!x$ as a modally strict theorem. By definition of $A!$, this implies $\exists x[\lambda x\, \neg\Diamond E!x]x$. Suppose a is such an object, so that we know $[\lambda x\neg\Diamond E!x]a$. Then $\neg\Diamond E!a$, by Rule $\overrightarrow{\beta}$C (184.1.a). By (162.1), $\Box\neg E!a$. Hence $\neg E!a$, by the T schema. Hence, by $\exists$I and $\exists$E, $\exists x\neg E!x$. By (103.2), the implies $\neg\forall xE!x$. Since this last result constitutes a modally strict theorem, it follows that $\Box\neg\forall xE!x$. ⋈

(228) Assume, for reductio, that $\exists x(O!x \,\&\, A!x)$. Suppose b is such an object, so that we know both $O!b$ and $A!b$. From the former, it follows that $\neg A!b$, by (222.2). Contradiction. ⋈

(229) If we let:

$$\varphi = O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)$$

$$\psi = O!x \,\&\, O!y \,\&\, x = y$$

then the following is an instance of axiom (49):

$$\begin{array}{l}[\lambda xy\, O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)]\downarrow \;\& \\ \quad \Box\forall x\forall y(O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy) \equiv O!x \,\&\, O!y \,\&\, x = y) \rightarrow \\ \qquad [\lambda xy\, O!x \,\&\, O!y \,\&\, x = y]\downarrow\end{array}$$

Since we want to show that the consequent of this axiom is a theorem, it suffices to establish the antecedent. But the first conjunct of the antecedent is axiomatic, since a core λ-expression is involved, by (39.2). So it remains only to establish the second conjunct of the antecedent. And by GEN and RN, it suffices to show:

$$O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy) \equiv O!x \,\&\, O!y \,\&\, x = y$$

($\rightarrow$) Assume $O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)$. Then it suffices to show $x = y$. But by $\vee$I, our assumption implies:

$$(O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)) \vee (A!x \,\&\, A!y \,\&\, \Box\forall F(xF \equiv yF))$$

Hence, $x = y$, by definition (23.1).

($\leftarrow$) Assume $O!x \,\&\, O!y \,\&\, x = y$. Then it suffices to show $\Box\forall F(Fx \equiv Fy)$. Note independently that $\Box\forall F(Fx \equiv Fx)$ is an easy modally strict theorem. From this and the 3rd conjunct of our assumption, it follows that $\Box\forall F(Fx \equiv Fy)$, by Rule =E. ⋈

(233.1) – **(233.2)** (Exercises)

(233.3) We established in *proof* of (229) that:

$$(O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)) \equiv (O!x \,\&\, O!y \,\&\, x = y)$$

and we established in (233.1) that:

$$x =_E y \equiv (O!x \,\&\, O!y \,\&\, x = y)$$

Then our theorem follows by simple biconditional reasoning from these two claims. ⋈

(234.1) ($\rightarrow$) Assume $x =_E y$. Then, by theorem (233.1), we know the following:

(ϑ) $O!x \,\&\, O!y \,\&\, x = y$

By (180.1), the first two conjuncts of (ϑ) imply, respectively, $\Box O!x$ and $\Box O!y$. The third conjunct of (ϑ) implies $\Box x = y$, by (125.1). Assembling what we have established:

$$\Box O!x \,\&\, \Box O!y \,\&\, \Box x = y$$

We now leave it as an exercise to show that we can extend theorem (158.3) to establish that $\Box\varphi \,\&\, \Box\psi \,\&\, \Box\chi$ implies $\Box(\varphi \,\&\, \psi \,\&\, \chi)$. Hence it follows that:

$$\Box(O!x \,\&\, O!y \,\&\, x = y)$$

But, independently, it follows from (233.1) by Rule RE that:

$$\Box x =_E y \equiv \Box(O!x \,\&\, O!y \,\&\, x = y)$$

Hence, our last two results imply $\Box x =_E y$, by (89.3.b). ($\leftarrow$) This direction is an instance of the T schema. ⋈

(234.2) ($\rightarrow$) From (234.1), it follows *a fortiori* that $x =_E y \rightarrow \Box x =_E y$. Since this is a $\Box$-theorem, it follows by (166.2) that $\Diamond x =_E y \rightarrow x =_E y$. ($\leftarrow$) This direction is an instance of the T$\Diamond$ schema. ⋈

(234.3) (Exercise)

(236) Observe that $[\lambda xy\, \neg(=_E xy)]\!\downarrow$ is a core λ-expression, notwithstanding the fact that $=_E$ is not a primitive constant; for every λ-expression of the form $[\lambda xy\, \neg\Pi xy]$, where Π is any binary relation term in which x and y don't occur free, is a core λ-expression, by definition (9.2). So as an instance of axiom (39.2), we know:

(ϑ) $[\lambda xy\, \neg(=_E xy)]\!\downarrow$

Then we may reason as follows:

$x \neq_E y$	$\equiv$	$\neq_E xy$	by convention (235.2)
	$\equiv$	$\overline{=_E}xy$	by convention (235.1)
	$\equiv$	$[\lambda xy\, \neg(=_E xy)]xy$	by Df (196), Rules $=_{df}$E and $=_{df}$I
	$\equiv$	$\neg(=_E xy)$	by (ϑ) and β-Conversion (48.2)
	$\equiv$	$\neg(x =_E y)$	by convention (231) ⋈

(237.1) ($\rightarrow$) The left-to-right direction of (234.1) is modally strict and so by RN, $\Box(x =_E y \rightarrow \Box x =_E y)$. So by theorem (171.1) $\Box(\neg x =_E y \rightarrow \Box\neg x =_E y)$, and so by the T schema, $\neg x =_E y \rightarrow \Box\neg x =_E y)$. Then by (236) and a Rule of Substitution, $x \neq_E y \rightarrow \Box x \neq_E y)$. ($\rightarrow$) By the T schema. ⋈

(237.1) [Alternative Proof:] It is a theorem of modal negation (162.1) that $\Box\neg x =_E y \equiv \neg\Diamond x =_E y$. Independently, we may negate both sides of (234.2) to conclude $\neg\Diamond x =_E y \equiv \neg x =_E y$. So by biconditional syllogism:

$$\Box\neg x =_E y \equiv \neg x =_E y$$

From this and the $\Box$-theorem derivable by commuting theorem (236), namely $\neg x =_E y \equiv x \neq_E y$, it follows by the Rule of Substitution (160.2) that $\Box x \neq_E y \equiv x \neq_E y$. Hence, $x \neq_E y \equiv \Box x \neq_E y$. ⋈

(237.2) ($\rightarrow$) It follows *a fortiori* from (237.1) that $x \neq_E y \rightarrow \Box x \neq_E y$. Since this is a $\Box$-theorem, it follows by (166.2) that $\Diamond x \neq_E y \rightarrow x \neq_E y$. ($\leftarrow$) This is an instance of the T$\Diamond$ schema. ⋈

(237.3) (Exercise)

(238.1) By the left-to-right direction of (234.1), it is a modally strict theorem that $x =_E y \rightarrow \Box x =_E y$. So by Rule RN, $\Box(x =_E y \rightarrow \Box x =_E y)$. Hence, by (174.2), $\mathcal{A}x =_E y \equiv x =_E y$. So $x =_E y \equiv \mathcal{A}x =_E y$. ⋈

(238.2) We may reason as follows:

$x \neq_E y$	$\equiv$	$\neg(x =_E y)$	by (236)
	$\equiv$	$\neg\mathcal{A}x =_E y$	by (238.1) and (88.4.b)
	$\equiv$	$\mathcal{A}\neg x =_E y$	by (44.1)
	$\equiv$	$\mathcal{A}x \neq_E y$	commute (236) and Rule (159.3) ⋈

(239.1) Assume $O!x$. So by (85.6), it follows that $O!x \,\&\, O!x$. Independently, by (117.1), we know $x = x$. So we have established: $O!x \,\&\, O!x \,\&\, x = y$. So by theorem (233.1), it follows that $x =_E x$. ⋈

(239.2) Assume $x =_E y$. Then, by (233.1) and &E, it follows *a fortiori* that $O!x$. Hence we know $x =_E x$, by (239.1). But it also follows from our assumption, by (233.2), that $x = y$. From these last two results, we may infer $y =_E x$, by Rule =E. ⋈

(239.3) Let φ be $y =_E z$ and let φ' be $x =_E z$. Then, the following is an instance of the axiom for the substitution of identicals (41):

$$(\xi)\quad y =_E x \rightarrow (y =_E z \rightarrow x =_E z)$$

But by (239.2), we know $x =_E y \rightarrow y =_E x$. From this and (ξ) it follows by hypothetical syllogism (76.3) that:

$$x =_E y \rightarrow (y =_E z \rightarrow x =_E z)$$

So by Importation (88.7.b), it follows that:

$$x =_E y \,\&\, y =_E z \rightarrow x =_E z \qquad ⋈$$

(240.1) Assume $O!x \vee O!y$. We reason by cases from the disjuncts.

Case 1. $O!x$. Our proof strategy is to show:

(ϑ) $O!x \vdash \Box O!x$

(ξ) $\Box O!x \vdash \Box(x = y \equiv x =_E y)$

For these two facts establish that $\Box(x = y \equiv x =_E y)$ is derivable from $O!x$, by (63.8).

Now (ϑ) follows from (180.1), by (63.10). To show (ξ), it suffices, by Rule RN, to show that there is a modally strict derivation of $x = y \equiv x =_E y$ from $O!x$, i.e., to show $O!x \vdash_\Box x = y \equiv x =_E y$. So, by (63.10), it suffices to show. $O!x \to (x = y \equiv x =_E y)$. $O!x$ is already an assumption of the present case. ($\to$) Assume $x = y$. From $O!x$ and (239.1), we know that $x =_E x$. Hence, by substitution of identicals, $x =_E y$. ($\leftarrow$) Assume $x =_E y$. Then by (233.2), $x = y$.

Case 2. By analogous reasoning. ⋈

(240.2) Assume $O!y$. Note independently that our conclusion is the consequent of the following instance of axiom (49):

$$([\lambda x\, x =_E y]{\downarrow} \;\&\; \Box\forall x(x =_E y \equiv x = y)) \to [\lambda x\, x = y]{\downarrow}$$

So it suffices to show:

(ϑ) $[\lambda x\, x =_E y]{\downarrow}$

(ξ) $\Box\forall x(x =_E y \equiv x = y)$

But (ϑ) is axiomatic (39.2), since $[\lambda x\, x =_E y]$ is a core λ-expression. To show (ξ), note that our assumption $O!y$ implies $O!x \vee O!y$, and so it follows from (240.1) that $\Box(x = y \equiv x =_E y)$. Since it is a modally strict theorem that biconditionals commute, it follows by a Rule of Substitution that $\Box(x =_E y \equiv x = y)$. Since x isn't free in our assumption, it follows by GEN that $\forall x \Box(x =_E y \equiv x = y)$. So by the Barcan Formula (167.1), it follows that $\Box\forall x(x = y \equiv x =_E y)$. ⋈

(241) Assume $\forall F(Fx \equiv Fy)$. Now consider the λ-expression $[\lambda x\, \Box\forall F(Fx \equiv Fy)]$. Since this is a core λ-expression, (39.2) asserts that it is significant. Moreover, $[\lambda x\, \Box\forall F(Fx \equiv Fy)]$ is substitutable for F in the matrix $Fx \equiv Fy$ of our assumption – no variable gets captured if we substitute $[\lambda x\, \Box\forall F(Fx \equiv Fy)]$ for F in this biconditional. Hence, by $\forall$E:

$$[\lambda x\, \Box\forall F(Fx \equiv Fy)]x \equiv [\lambda x\, \Box\forall F(Fx \equiv Fy)]y$$

But we also know that β-Conversion applies to both sides of this biconditional, i.e., that:

$$[\lambda x\, \Box\forall F(Fx \equiv Fy)]x \equiv \Box\forall F(Fx \equiv Fy)$$

$$[\lambda x\, \Box\forall F(Fx \equiv Fy)]y \equiv \Box\forall F(Fy \equiv Fy)$$

Hence it follows that $\Box\forall F(Fx \equiv Fy) \equiv \Box\forall F(Fy \equiv Fy)$. But the right side of this last result is a theorem. Hence $\Box\forall F(Fx \equiv Fy)$. ⋈

(242.1) Assume $O!x$ and $\forall F(Fx \equiv Fy)$. Then $O!y$. Moreover, $\forall F(Fx \equiv Fy)$ implies $\Box\forall F(Fx \equiv Fy)$, by the theorem (241). So by the right-to-left direction of (233.3), it follows from our first two assumptions and this last result that $x =_E y$. ⋈

(242.2) (Exercise)

(243.1) Assume $O!x$, $O!y$. Note that both $[\lambda z\, z =_E x]\downarrow$ and $[\lambda z\, z =_E y]\downarrow$, by now familiar reasoning. $(\rightarrow)$ Assume $x \neq y$. Then $x \neq_E y$, by (233.2). For reductio, assume $[\lambda z\, z =_E x] = [\lambda z\, z =_E y]$. Since $O!x$, we know by the reflexivity of $=_E$ (239.1) that $x =_E x$. Since $[\lambda z\, z =_E x]\downarrow$ and $x\downarrow$, it follows by Rule $\overleftarrow{\beta}$C (184.2.a) that $[\lambda z\, z =_E x]x$. But then by Rule =E, it follows that $[\lambda z\, z =_E y]x$. By Rule $\overrightarrow{\beta}$C (184.1.a), it follows that $x =_E y$. But by (236), this contradicts $x \neq_E y$. $(\leftarrow)$ The proof of this direction, which is equivalent to its contrapositive $x = y \rightarrow [\lambda z\, z =_E x] = [\lambda z\, z =_E y]$, is trivial.[455] ⋈

(243.2) (Exercise)

(244) Assume $O!x$. Then by (180.1), it follows that $\Box O!x$. Since the closures of the instances of (52) are axioms, we also know: $\Box(O!x \rightarrow \neg\exists FxF)$. So by the K axiom (45.1), $\Box O!x \rightarrow \Box\neg\exists FxF$ and, hence, $\Box\neg\exists FxF$. ⋈

(245.1) Assume $\forall F(xF \equiv yF)$. By the Barcan formula (167.1), it suffices to show, $\forall F\Box(xF \equiv yF)$. Since F isn't free in our assumption, it suffices by GEN to show $\Box(xF \equiv yF)$. By (179.5) (set G in (179.5) to F), it suffices to show $\Box xF \equiv \Box yF$. But we may establish this by the following biconditional chain:

$$\begin{array}{lll} \Box xF & \equiv & xF \quad \text{by (179.2)} \\ & \equiv & yF \quad \text{by our assumption} \\ & \equiv & \Box yF \quad \text{by (179.2)} \quad ⋈ \end{array}$$

(245.2) Suppose $A!x$, $A!y$, and $\forall F(xF \equiv yF)$. The third assumption implies, by (245.1), that $\Box\forall F(xF \equiv yF)$. From this last fact and our first two assumptions, it follows by (23.1) that $x = y$. ⋈

(245.3) (Exercise)

(246) By contraposing (52) and eliminating the double negation, we have: $\exists FxF \rightarrow \neg O!x$. But the right-to-left direction of (222.3) is $\neg O!x \rightarrow A!x$. So by hypothetical syllogism, $\exists FxF \rightarrow A!x$ ⋈

[455] In an earlier draft of this work, the statement of the contrapositive in the proof was inadvertently reversed. The error was discovered by Daniel Kirchner, using automated reasoning techniques. See Kirchner 2017 [2021] and 2022.

(247.1) By axiom (39.2), $H\downarrow$. So by the unary case of definition (20.2), $\exists x\, xH$. By GEN, $\forall H \exists x\, xH$. ⋈

(247.2) – **(247.4)** (Exercises)

(250) Let φ be any formula with no free xs. Then by (53), $\exists x(A!x \,\&\, \forall F(xF \equiv \varphi))$. Assume a is an arbitrary such object, so that we know:

(ϑ) $A!a \,\&\, \forall F(aF \equiv \varphi)$

To show $\exists! x(A!x \,\&\, \forall F(xF \equiv \varphi))$, it suffices, by &I, $\exists$I, the definition of the uniqueness quantifier (127.1) and $\exists$E, to show:

$$\forall y[(A!y \,\&\, \forall F(yF \equiv \varphi)) \to y = a]$$

But by GEN, it then suffices to show $A!y \,\&\, \forall F(yF \equiv \varphi) \to y = a$. So assume:

(ξ) $A!y \,\&\, \forall F(yF \equiv \varphi)$

The second conjuncts of (ϑ) and (ξ) jointly imply $\forall F(aF \equiv yF)$, by the laws of quantified biconditionals (99.11) and (99.10). It then follows from this, by (245.2), that $a = y$, since we already know $A!a$ and $A!y$ as the first conjuncts of (ϑ) and (ξ). So $y = a$, by the symmetry of identity. ⋈

(251.1) – **(251.6)** These are all instances of theorem (250). ⋈

(252) Let φ be any formula in which x doesn't occur free. Then by Strengthened Comprehension (250), it is a theorem that $\exists! x(A!x \,\&\, \forall F(xF \equiv \varphi))$. Hence, by the Rule of Actualization (135), it follows that $\mathcal{A}\exists! x(A!x \,\&\, \forall F(xF \equiv \varphi))$. So by (176.2), $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))\downarrow$. ⋈

(254)★ (Exercise)

(255) Consider any formula φ in which x doesn't occur free. Now let ψ be the formula $A!x \,\&\, \forall F(xF \equiv \varphi)$. Then if we assume the antecedent of our theorem, our assumption can be expressed as:

(ϑ) $y = \imath x\psi$

Without loss of generality, choose z be to some variable substitutable for x in φ. Then, independently, as an instance of the modally strict version of Hintikka's schema (148), we know:

$$y = \imath x\psi \to (\mathcal{A}\psi^y_x \,\&\, \forall z(\mathcal{A}\varphi^z_x \to z = y))$$

From the last fact and (ϑ), it follows that:

$$\mathcal{A}\psi^y_x \,\&\, \forall z(\mathcal{A}\varphi^z_x \to z = y)$$

i.e.,

$$\mathcal{A}(A!y \,\&\, \forall F(yF \equiv \varphi)) \,\&\, \forall z(\mathcal{A}(A!z \,\&\, \forall F(zF \equiv \varphi)) \to z = y)$$

By (139.2), $\mathcal{A}$ distibutes over a conjunction and so the first conjunct of our last result implies:

$$\mathcal{A}A!y \,\&\, \mathcal{A}\forall F(yF \equiv \varphi)$$

The first conjunct of this result, namely $\mathcal{A}A!y$, implies $A!y$, by the modally-strict theorem (180.8). ⋈

(256.1)★ Suppose φ is a formula in which x doesn't occur free. Then by (252), we know that $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))\downarrow$. We leave it as an exercise to show that the description $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$ is substitutable for x in its own matrix. Hence by (145.3)★, we may perform the substitution to conclude:

$$A!\imath x(A!x \,\&\, \forall F(xF \equiv \varphi)) \,\&\, \forall F(\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))F \equiv \varphi)$$

The second conjunct of this result is:

$$\forall F(\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))F \equiv \varphi)$$

So by the special case of Rule ∀E (93.3):

$$\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))F \equiv \varphi \qquad ⋈$$

(256.2)★ In the proof of (256.1)★ we established:

$$\forall F(\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))F \equiv \varphi)$$

Since G is, by hypothesis, substitutable for F in φ, it is substitutable for F in the matrix of the above universal claim. And since $G\downarrow$, we may use Rule ∀E (93.1) to instantiate G for $\forall F$ in the above to obtain:

$$\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G \equiv \varphi^G_F \qquad ⋈$$

(258.1) Consider any formula φ where x doesn't occur free. We want to show:

$$\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))F \equiv \mathcal{A}\varphi$$

In what follows, we take ψ to be: $A!x \,\&\, \forall F(xF \equiv \varphi)$. Hence, we want to show:

$$\imath x\psi F \equiv \mathcal{A}\varphi$$

Now $\imath x\psi$ is a canonical description, and so we know, by (252), that $\imath x\psi\downarrow$. So by (152.4), it follows that $\mathcal{A}\psi^{\imath x\psi}_x$, i.e.,

$$\mathcal{A}(A!\imath x\psi \,\&\, \forall F(\imath x\psi F \equiv \varphi))$$

Since the actuality operator distributes over a conjunction (139.2), it follows that $\mathcal{A}\forall F(\imath x\psi F \equiv \varphi)$. Since the actuality operator and the universal quantifier commute (44.3), we may infer $\forall F\mathcal{A}(\imath x\psi F \equiv \varphi)$. Hence, by the special case of Rule $\forall$E, $\mathcal{A}(\imath x\psi F \equiv \varphi)$. So by (139.5), $\mathcal{A}\imath x\psi F \equiv \mathcal{A}\varphi$. But we also know, by the commuted version of (179.10) and the fact that $(\imath x\psi)\downarrow$, that $\imath x\psi F \equiv \mathcal{A}\imath x\psi F$. So by biconditional syllogism, $\imath x\psi F \equiv \mathcal{A}\varphi$, which is what we had to show ⋈

(258.2) (Exercise)

(259.1) Assume $\Box\varphi^G_F$, where x doesn't occur free in φ and G is substitutable for F in φ. It follows that $\mathcal{A}\varphi^G_F$, by theorem (132). But then by the right-to-left direction of the theorem (258.2), we may conclude $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G$. ⋈

(259.2) Assume:

(ϑ) $\Box\varphi^G_F$

where φ is any formula in which both x doesn't occur free in φ and G is substitutable for F. Then by a 'paradox' of strict implication (158.1), (ϑ) implies:

(ξ) $\Box(\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G \rightarrow \varphi^G_F)$

Put this result aside for the moment. By our previous theorem (259.1), it also follows from (ϑ) that:

$$\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G$$

From this and the fact that every canonically-described individual exists (252) it follows by the rigidity of encoding that:

$$\Box\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G$$

This implies, by the same 'paradox' of strict implication, that:

(ζ) $\Box(\varphi^G_F \rightarrow \imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G)$

By &I we may conjoin (ξ) and (ζ), so that by (158.4), we've derived:

$$\Box(\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G \equiv \varphi^G_F)$$ ⋈

(261.1) Suppose φ is a rigid condition on properties (260.1) in which x doesn't occur free. Now assume the antecedent of our theorem, so that we know both:

(a) $A!x$

(b) $\forall F(xF \equiv \varphi)$

To show $\Box(A!x \& \forall F(xF \equiv \varphi))$, it suffices, by the right-to-left direction of (158.3), to show both $\Box A!x$ and $\Box\forall F(xF \equiv \varphi)$. But $\Box A!x$ follows from (a) by (180.2). So it remains to show $\Box\forall F(xF \equiv \varphi)$. By BF (167.1), it suffices to show $\forall F\Box(xF \equiv \varphi)$ and by GEN, it suffices to show $\Box(xF \equiv \varphi)$. But note that this is the consequent of the following instance of (172.4):

$$(\Box(\varphi \to \Box\varphi) \,\&\, \Box(xF \to \Box xF)) \to ((\Box xF \equiv \Box\varphi) \to \Box(xF \equiv \varphi))$$

So it remains to establish:

(ϑ) $\Box(\varphi \to \Box\varphi)$

(ξ) $\Box(xF \to \Box xF)$

(ζ) $\Box xF \equiv \Box\varphi$

For (ϑ), note that since φ is, by hypothesis, a rigid condition on properties, it follows by definition (260.1) that $\vdash_{\Box} \forall F(\varphi \to \Box\varphi)$. Hence $\vdash_{\Box} \varphi \to \Box\varphi$, by Rule $\forall$E (93.3). So by RN, $\Box(\varphi \to \Box\varphi)$ is a theorem. ⋈

For (ξ), we simply note that it is an axiom; it is a closure of (51).

For (ζ), we prove both directions. ($\to$) Assume $\Box xF$. Then xF by the T schema (45.2). But (b) implies $xF \equiv \varphi$. Hence, φ. Moreover, by the T schema, (ϑ) implies $\varphi \to \Box\varphi$. Hence $\Box\varphi$. ($\leftarrow$) Assume $\Box\varphi$. Then φ, by the T schema. Since (b) implies $xF \equiv \varphi$, it follows that xF, which implies $\Box xF$ by axiom (51). ⋈

(261.2) Suppose φ is a rigid condition on properties in which x doesn't occur free. Now let ψ be the formula $A!x \,\&\, \forall F(xF \equiv \varphi)$. Then it follows from (261.1) by GEN that:

$$\forall x(\psi \to \Box\psi)$$

So by theorem (153.2), it follows that:[456]

$$\exists!x\psi \to \forall y(y = \iota x\psi \to \psi^{y}_{x})$$

But we know $\exists!x\psi$, since that just is Strenghtened Comprehension for Abstract Objects (250), i.e.,

$$\exists!x(A!x \,\&\, \forall F(xF \equiv \varphi))$$

So we may conclude $\forall y(y = \iota x\psi \to \psi^{y}_{x})$, i.e.,

$$\forall y(y = \iota x(A!x \,\&\, \forall F(xF \equiv \varphi)) \to (A!y \,\&\, \forall F(yF \equiv \varphi)))$$

Hence, our theorem follows by Rule $\forall$E (93.3). ⋈

(261.3) Let φ be a rigid condition on properties in which both x doesn't occur free. Now we know, independently, by (252), that:

$$\iota x(A!x \,\&\, \forall F(xF \equiv \varphi))\downarrow$$

[456] I'm indebted to Daniel Kirchner for discovering, using automated reasoning techniques, that an incorrect theorem was cited in a previous version of this proof. See Kirchner 2017 [2021] and 2022.

For ease of reading, abbreviate the description as $\imath x\psi$, so that we know $\imath x\psi\downarrow$. Then the description can be instantiated in the universal generalization of (261.2) to obtain:

$$\imath x\psi = \imath x\psi \rightarrow (A!\imath x\psi \,\&\, \forall F(\imath x\psi F \equiv \varphi))$$

Since the antecedent holds by Rule =I, we may detach the consequent. The second conjunct of the consequent is $\forall F(\imath x\psi F \equiv \varphi)$, i.e.,

$$\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))F \equiv \varphi \qquad \bowtie$$

(264.1) By definition (263.1) and the Rule of Substitution for Defined Formulas (160.3), we have to show:

(ϑ) $\exists!x(A!x \,\&\, \neg\exists FxF)$

To find a witness, note that by the Comprehension Principle for Abstract Objects (53), we know:

$$\exists x(A!x \,\&\, \forall F(xF \equiv F \neq F))$$

Suppose a is an arbitrary such object, so that we know:

(ξ) $A!a \,\&\, \forall F(aF \equiv F \neq F)$

Our strategy is to show that a is a witness to (ϑ). So by &I, $\exists$I, the definition of the uniqueness quantifier, and Rule $\equiv_{df}$I, it suffices to show both:

(a) $A!a \,\&\, \neg\exists FaF$

(b) $\forall y((A!y \,\&\, \neg\exists FyF) \rightarrow y = a)$

(a) Since the first conjunct of (ξ) is $A!a$, it remains to show $\neg\exists FaF$. Suppose, for reductio, that $\exists FaF$ and that P is an arbitrary such property, so that we know aP. Then by the right conjunct of (ξ), it follows that $P \neq P$, contradicting the fact that $P = P$, which we know by the special case of Rule =I (118.2).

(b) By GEN, it suffices to show $(A!y \,\&\, \neg\exists FyF) \rightarrow y = a$. So assume $A!y \,\&\, \neg\exists FyF$. But the second conjunct, $\neg\exists FyF$, and the fact we just established, namely $\neg\exists FaF$, jointly imply $\forall F(aF \equiv yF)$, by (103.9). Since we also know that both y and a are abstract, it follows by (245.2) that $y = a$. $\bowtie$

(264.2) (Exercise)

(264.3) Since (264.1) is a modally strict theorem, it follows that $\mathcal{A}\exists!xNull(x)$, by the Rule of Actualization. So by (176.2), $\imath xNull(x)\downarrow$. $\bowtie$

(264.4) (Exercise)

(266.1) Assume $Null(x)$. Then, by definition (263.1) and &E, we know both (a) $A!x$ and (b) $\neg\exists FxF$. Now to show $\Box Null(x)$, we have to show $\Box(A!x \& \neg\exists FxF)$, by the Rule of Substitution for Defined Formulas (160.3). By &I and (158.3), it suffices to show $\Box A!x$ and $\Box\neg\exists FxF$. By (180.2), (a) implies $\Box A!x$. Now to show $\Box\neg\exists FxF$, suppose $\neg\Box\neg\exists FxF$, for reductio. Then by definition (18.5) and Rule $\equiv_{df}$I, $\Diamond\exists FxF$. So by BF$\Diamond$ (167.3), $\exists F\Diamond xF$. Suppose P is an arbitrary such property, so that we know $\Diamond xP$. Then by the left-to-right direction of (179.3), it follows that xP, and hence, $\exists FxF$, by $\exists$I, which contradicts (b). ⋈

(266.2) Assume $Universal(x)$. Then, by definition (263.2), we know both (a) $A!x$ and (b) $\forall FxF$. Now to show $\Box Universal(x)$, we have to show $\Box(A!x \,\&\, \forall FxF)$, by the Rule of Substitution for Defined Formulas (160.3). By &I and (158.3), it suffices to show $\Box A!x$ and $\Box\forall FxF$. By (180.2), (a) implies $\Box A!x$. Now to show $\Box\forall FxF$, suppose $\neg\Box\forall FxF$, for reductio. Then by (158.11), $\Diamond\neg\forall FxF$. Since $\neg\forall FxF \equiv \exists F\neg xF$ is an instance of the modally strict theorem (103.2), it follows by the Rule of Substitution (160.2) that $\Diamond\exists F\neg xF$. So by BF$\Diamond$ (167.3), $\exists F\Diamond\neg xF$. Suppose P is an arbitrary such property, so that we know $\Diamond\neg xP$. Then by the left-to-right direction of (179.8), it follows that $\neg xP$, and hence, $\exists F\neg xF$, i.e., $\neg\forall FxF$, which contradicts (b). ⋈

(266.3) As an instance of (153.2), we know:

$$\forall x(Null(x) \to \Box Null(x)) \to (\exists!xNull(x) \to \forall y(y = \imath xNull(x) \to Null(y)))$$

By applying GEN to (266.1), we know $\forall x(Null(x) \to \Box Null(x))$. Hence:

$$\exists!xNull(x) \to \forall y(y = \imath xNull(x) \to Null(y))$$

But (264.1) is $\exists!xNull(x)$. Hence:

$$\forall y(y = \imath xNull(x) \to Null(y))$$

Since $\boldsymbol{a}_\varnothing\downarrow$, it follows that:

$$\boldsymbol{a}_\varnothing = \imath xNull(x) \to Null(\boldsymbol{a}_\varnothing)$$

But by definition (265.1) and Rule of Identity by Definition (120.1), $\boldsymbol{a}_\varnothing = \imath xNull(x)$. Hence $Null(\boldsymbol{a}_\varnothing)$. ⋈

(266.4) (Exercise)

(266.5) By (266.4) and (263.2), it follows that $A!\boldsymbol{a}_{\boldsymbol{V}}$ and $\forall F\boldsymbol{a}_{\boldsymbol{V}}F$. By (266.3) and (263.1), it follows that $A!\boldsymbol{a}_\varnothing$ and $\neg\exists F\boldsymbol{a}_\varnothing F$. Since both $\boldsymbol{a}_{\boldsymbol{V}}$ and $\boldsymbol{a}_\varnothing$ are abstract, it suffices, by (245.3), to show $\exists F(\boldsymbol{a}_{\boldsymbol{V}}F \& \neg\boldsymbol{a}_\varnothing F)$. Given that we can prove the existence of at least some properties, any property you please serves as a witness (exercise). ⋈

(266.6) – **(266.7)** (Exercises)

(268.1) [Note: Readers who have gotten this far should now be well-equipped to follow the more obvious shortcuts in reasoning we take in this proof and the ones that follow.] Consider the following instance of comprehension, in which there is a free occurrence of the binary relation variable F:

$$\exists x(A!x \,\&\, \forall G(xG \equiv \exists y(A!y \,\&\, G = [\lambda z\, Fzy] \,\&\, \neg yF)))$$

Assume a is an arbitrary such object, so that we have:

$$(\vartheta)\quad A!a \,\&\, \forall G(aG \equiv \exists y(A!y \,\&\, G = [\lambda z Fzy] \,\&\, \neg yF))$$

Now consider the property $[\lambda z\, Fza]$, which we know exists by (39.2), and we now show, by reductio, that a encodes this property. Assume $\neg a[\lambda z\, Fza]$. Then, from the second conjunct of (ϑ) it follows that:

$$\neg\exists y(A!y \,\&\, [\lambda z\, Fza] = [\lambda z\, Fzy] \,\&\, \neg y[\lambda z\, Fza]))),$$

i.e., by quantifier negation:

$$(\xi)\quad \forall y((A!y \,\&\, [\lambda z\, Fza] = [\lambda z\, Fzy]) \rightarrow y[\lambda z\, Fza]),$$

i.e., for any abstract object y, if $[\lambda z\, Fza] = [\lambda z\, Fzy]$, then $y[\lambda z\, Fza]$. Instantiate (ξ) to a. We know $A!a$ by the left conjunct of (ϑ) and we know $[\lambda zFza] = [\lambda zFza]$ by Rule =I. Hence $a[\lambda z\, Fza]$, contrary to assumption. So we've established by reductio that $a[\lambda z\, Fza]$. Then by the second conjunct of (ϑ), there is an abstract object, say b, such that both $[\lambda z\, Fza] = [\lambda z\, Fzb]$ and $\neg b[\lambda z\, Fza]$. But since $a[\lambda z\, Fza]$ and $\neg b[\lambda z\, Fza]$, it follows by the contrapositive of Rule =E that $a \neq b$.[457] So, by two sequences of applying $\exists$I and then $\exists$E (i.e., first generalizing on, and then discharging b, and then generalizing on, and then discharging a), we've established that there are abstract objects x and y such that $x \neq y$, yet such that $[\lambda z\, Fzx] = [\lambda z\, Fzy]$. By GEN, this theorem holds for every binary relation F. ⋈

(268.2) By reasoning analogous to that used in the proof of (268.1). ⋈

(268.3) Consider the following instance of comprehension in which there is a free occurrence of the property variable F:

$$\exists x(A!x \,\&\, \forall G(xG \equiv \exists y(A!y \,\&\, G = [\lambda z\, Fy] \,\&\, \neg yG)))$$

By reasoning analogous to that used in the proof of (268.1), it is straightforward to establish that there are distinct abstract objects, say k and l, such that $[\lambda z\, Fk]$ is identical to $[\lambda z\, Fl]$. But, then by the fundamental theorem governing proposition identity (116.3), it follows that $Fk = Fl$. We know generally, by (111.1) and symmetry of identity, that $\varphi = [\lambda\, \varphi]$. So it follows from $Fk = Fl$

[457] Since Rule =E is: φ^{τ}_{α}, $\tau = \tau' \vdash \varphi'$, its contrapositive rule is: φ^{τ}_{α}, $\neg\varphi' \vdash \tau \neq \tau'$. We leave its justification as an exercise.

that $[\lambda\, Fk] = [\lambda\, Fl]$. Hence, there are distinct abstract objects x and y such that $[\lambda\, Fx] = [\lambda\, Fy]$. By GEN, this holds for every property F. ⋈

(269) Let R_1 be the relation $[\lambda xy\, \forall F(Fx \equiv Fy)]$, which exists by (39.2). By (268.1), it follows that:

$$\exists x \exists y(A!x \,\&\, A!y \,\&\, x \neq y \,\&\, [\lambda z\, R_1 zx] = [\lambda z\, R_1 zy])$$

Suppose a and b are such objects, so that we know:

(ϑ) $A!a \,\&\, A!b \,\&\, a \neq b \,\&\, [\lambda z\, R_1 za] = [\lambda z\, R_1 zb]$

Since we know both $R_1\!\downarrow$ and $a\!\downarrow$, we can apply Rule $\overleftarrow{\beta}$C (184.2.a) to the easily-established fact that $\forall F(Fa \equiv Fa)$ and conclude $R_1 aa$. Note independently that by (39.2), $[\lambda z\, [\lambda xy\, \forall F(Fx \equiv Fy)]za]\!\downarrow$, i.e., $[\lambda z\, R_1 za]\!\downarrow$. Hence by Rule $\overleftarrow{\beta}C$ (184.2.a) we may infer from the previously established $R_1 aa$ that $[\lambda z\, R_1 za]a$. From this and the fourth conjunct of (ϑ), it follows by Rule =E that $[\lambda z\, R_1 zb]a$. Hence, by Rule $\overrightarrow{\beta}$C (184.1.a), $R_1 ab$. Since R_1 is $[\lambda xy\, \forall F(Fx \equiv Fy)]$, we may apply Rule $\overrightarrow{\beta}$C again to conclude $\forall F(Fa \equiv Fb)$. Hence, $\exists x \exists y(A!x \,\&\, A!y \,\&\, a \neq b \,\&\, \forall F(Fx \equiv Fy))$. ⋈

(271.1) Let φ be any formula in which y doesn't occur free. (→) Our proof strategy is:

(A) Assume $[\lambda x\, \varphi]\!\downarrow$, for conditional proof

(B) Infer $\Box[\lambda x\, \varphi]\!\downarrow$, from (A) by (106)

(C) Establish: $\Box[\lambda x\, \varphi]\!\downarrow \to \Box\forall x \forall y(\forall F(Fx \equiv Fy) \to (\varphi \equiv \varphi^y_x))$

(D) Conclude $\Box\forall x \forall y(\forall F(Fx \equiv Fy) \to (\varphi \equiv \varphi^y_x))$, from (B) and (C)

Given this strategy, we need only prove (C). By Rule RM, it suffices to show that there is a modally strict proof of:

$$[\lambda x\, \varphi]\!\downarrow \to \forall x \forall y(\forall F(Fx \equiv Fy) \to (\varphi \equiv \varphi^y_x))$$

So assume $[\lambda x\, \varphi]\!\downarrow$. Since x and y don't occur free in any assumption, it suffices by GEN to show $\forall F(Fx \equiv Fy) \to (\varphi \equiv \varphi^y_x)$. So assume $\forall F(Fx \equiv Fy)$. Since $[\lambda x \varphi]\!\downarrow$ and is substitutable for F in the matrix of our last assumption, it follows that $[\lambda x \varphi]x \equiv [\lambda x \varphi]y$. But since we can apply β-Conversion to both sides, it follows that $\varphi \equiv \varphi^y_x$. ⋈

(←) Our proof strategy is:

(E) Assume: $\Box\forall x \forall y(\forall F(Fx \equiv Fy) \to \varphi \equiv \varphi^y_x)$, for conditional proof.

(F) Show: $\Box\forall x \forall y(\forall F(Fx \equiv Fy) \to (\varphi \equiv \varphi^y_x)) \to \Box\forall y(\exists x(\forall F(Fx \equiv Fy) \,\&\, \varphi) \equiv \varphi^y_x)$.

(G) Infer: $\Box\forall y(\exists x(\forall F(Fx \equiv Fy) \,\&\, \varphi) \equiv \varphi^y_x)$, from (E) and (F).

(H) Infer: $[\lambda y\, \varphi^y_x]\downarrow$, by appealing to the following alphabetic-variant of axiom (49):

$$[\lambda y\, \exists x(\forall F(Fx \equiv Fy) \,\&\, \varphi)]\downarrow \,\&\, \Box\forall y(\exists x(\forall F(Fx \equiv Fy) \,\&\, \varphi) \equiv \varphi^y_x) \to [\lambda y\, \varphi^y_x]\downarrow$$

and deriving both conjuncts of the antecedent. The first conjunct of the antecedent of the above, namely $[\lambda y \exists x(\forall F(Fx \equiv Fy) \& \varphi)]\downarrow$, is an instance of (39.2); since y doesn't occur free in φ by hypothesis, the λ-expression is a core λ-expression (9.2) – no variable bound by the λ occurs in encoding position (9.1) in the matrix. The second conjunct of the antecedent of the instance of (49) is just (G).

(I) Conclude: $[\lambda x\, \varphi]\downarrow$, since this is just an alphabetic variant of (H).

Given this proof strategy, it remains to show (F). By Rule RM, it suffices to show that there is a modally strict proof of:

$$\forall x\forall y(\forall F(Fx \equiv Fy) \to (\varphi \equiv \varphi^y_x)) \to \forall y(\exists x(\forall F(Fx \equiv Fy) \,\&\, \varphi) \equiv \varphi^y_x)$$

So assume $\forall x\forall y(\forall F(Fx \equiv Fy) \to (\varphi \equiv \varphi^y_x))$. Since y doesn't occur free in our assumption, it suffices by GEN to show: $\exists x(\forall F(Fx \equiv Fy) \,\&\, \varphi) \equiv \varphi^y_x$. So, for reductio, assume not. Then either (J) or (K):

(J) $\exists x(\forall F(Fx \equiv Fy) \,\&\, \varphi) \,\&\, \neg\varphi^y_x$

(K) $\varphi^y_x \,\&\, \neg\exists x(\forall F(Fx \equiv Fy) \,\&\, \varphi)$

But both lead to contradiction.

If (J), then suppose a is a witness to the first conjunct, so that we know:

(L) $\forall F(Fa \equiv Fy) \,\&\, \varphi^a_x$

If we instantiate a and y, respectively, into our initial assumption, it follows that $\forall F(Fa \equiv Fy) \to (\varphi^a_x \equiv \varphi^y_x)$. This and the first conjunct of (L) imply $\varphi^a_x \equiv \varphi^y_x$. But this result and the second conjunct of (L) imply φ^y_x, which contradicts the second conjunct of (J).

If (K), then the second conjunct implies:

(M) $\forall x(\neg\forall F(Fx \equiv Fy) \vee \neg\varphi)$

If we instantiate (M) to y, then either $\neg\forall F(Fy \equiv Fy)$ or $\neg\varphi^y_x$. Since $\forall F(Fy \equiv Fy)$ is a theorem, it follows that $\neg\varphi^y_x$, which contradicts the first conjunct of (K). ⋈

(271.2) – (271.3) (Exercises)

(272.1) Let φ be any formula in which y doesn't occur free. Assume $[\lambda x\, \varphi]\downarrow$. Then by (271.1), $\Box\forall x\forall y(\forall F(Fx \equiv Fy) \to (\varphi \equiv \varphi^y_x))$. By now familiar modal reasoning using the Converse Barcan Formula and a Rule of Substitution, it follows that:

(ϑ) $\forall x \forall y \Box(\forall F(Fx \equiv Fy) \to (\varphi \equiv \varphi^y_x))$

Now we want to show $\forall x \forall y(\forall F(Fx \equiv Fy) \to \Box(\varphi \equiv \varphi^y_x))$. Since x and y don't occur free in any assumptions we've made, it suffices by GEN to show $\forall F(Fx \equiv Fy) \to \Box(\varphi \equiv \varphi^y_x)$. So assume $\forall F(Fx \equiv Fy)$. Then if we instantiate x and y into (ϑ), it follows that:

$$\Box(\forall F(Fx \equiv Fy) \to (\varphi \equiv \varphi^y_x))$$

By the K axiom, this implies:

(ξ) $\Box \forall F(Fx \equiv Fy) \to \Box(\varphi \equiv \varphi^y_x)$

But it follows from our assumption $\forall F(Fx \equiv Fy)$ that $\Box \forall F(Fx \equiv Fy)$, by (241). Hence $\Box(\varphi \equiv \varphi^y_x)$. ⋈

(272.2) (Exercise)

(272.3) [Contributed by Daniel Kirchner] Let w, x, y, z all be unrestricted individual variables. Then note that the term $[\lambda x\, w[\lambda z Gxz]]$ is a core λ-expression. So $[\lambda x\, w[\lambda z\, Gxz]]\downarrow$ is an instance of axiom (39.2). Now consider the following instance of the Corollary to Kirchner Theorem (272.1):

$$[\lambda x\, w[\lambda z\, Gxz]]\downarrow \to \forall x \forall y(\forall F(Fx \equiv Fy) \to \Box(w[\lambda z\, Gxz] \equiv w[\lambda z\, Gyz]))$$

Since the antecedent is axiomatic, we may infer:

$$\forall x \forall y(\forall F(Fx \equiv Fy) \to \Box(w[\lambda z\, Gxz] \equiv w[\lambda z\, Gyz]))$$

And for an arbitrary x and y, this reduces by the special case of $\forall$E to:

$$\forall F(Fx \equiv Fy) \to \Box(w[\lambda z\, Gxz] \equiv w[\lambda z\, Gyz])$$

Since we derived the above from no assumptions and w is free, it follows by GEN:

$$\forall w(\forall F(Fx \equiv Fy) \to \Box(w[\lambda z\, Gxz] \equiv w[\lambda z\, Gyz]))$$

By theorem (95.2), the above implies:

$$\forall F(Fx \equiv Fy) \to \forall w \Box(w[\lambda z\, Gxz] \equiv w[\lambda z\, Gyz])$$

And since the Barcan formulas (167.1) and (167.2) yield the modally strict equivalence $\forall \alpha \Box \varphi \equiv \Box \forall \alpha \varphi$, it follows by a Rule of Substitution that:

(ϑ) $\forall F(Fx \equiv Fy) \to \Box \forall w(w[\lambda z\, Gxz] \equiv w[\lambda z\, Gyz])$

Now if we can show:

(ζ) $\Box \forall w(w[\lambda z\, Gxz] \equiv w[\lambda z\, Gyz]) \to [\lambda z\, Gxz] = [\lambda z\, Gyz]$

then our theorem:

$$\forall F(Fx \equiv Fy) \rightarrow [\lambda z\, Gxz] = [\lambda z\, Gyz]$$

will follow by a simple hypothetical syllogism from (ϑ) and (ζ). So it remains to show (ζ). Assume $\Box\forall w(w[\lambda z\, Gxz] \equiv w[\lambda z\, Gyz])$, to show $[\lambda z\, Gxz] = [\lambda z\, Gyz]$. Note that by definition (23.2) and the Rule of Biconditional Simplification, the following derivation is valid:

$$[\lambda z\, Gxz]\downarrow, [\lambda z\, Gyz]\downarrow \vdash_{\Box} [\lambda z\, Gxz] = [\lambda z\, Gyz] \equiv \Box\forall w(w[\lambda z\, Gxz] \equiv w[\lambda z\, Gyz])$$

But the two premises of this derivation are both axioms, since the terms asserted to be significant are core λ-expressions. Hence it follows that:

$$[\lambda z\, Gxz] = [\lambda z\, Gyz] \equiv \Box\forall w(w[\lambda z\, Gxz] \equiv w[\lambda z\, Gyz])$$

But since we know the right-hand side by assumption, it follows that $[\lambda z Gxz] = [\lambda z\, Gyz]$. ⋈

(272.4) (Exercise)

(272.5) Let w, x, y, z all be unrestricted individual variables. Assume $[\lambda z\, \varphi]\downarrow$, where none of the free occurrences of x in φ are in encoding position (9.1). Note that this condition implies that the term $[\lambda x\, w[\lambda z \varphi]]$ is a core λ-expression (9.2). So $[\lambda x\, w[\lambda z\, \varphi]]\downarrow$ is an instance of axiom (39.2). Now consider the following instance of the Corollary to Kirchner Theorem (272.1):

$$[\lambda x\, w[\lambda z\, \varphi]]\downarrow \rightarrow \forall x \forall y(\forall F(Fx \equiv Fy) \rightarrow \Box(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi^y_x]))$$

Since the antecedent is axiomatic, we may infer:

$$\forall x \forall y(\forall F(Fx \equiv Fy) \rightarrow \Box(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi^y_x]))$$

So for an arbitrary x and y, it follows that:

$$\forall F(Fx \equiv Fy) \rightarrow \Box(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi^y_x])$$

Since we derived the above from no assumptions with w free, it follows by GEN:

$$\forall w(\forall F(Fx \equiv Fy) \rightarrow \Box(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi^y_x]))$$

So by (95.2):

$$\forall F(Fx \equiv Fy) \rightarrow \forall w \Box(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi^y_x])$$

By a Rule of Substitution and the modally strict equivalence of the Barcan Formulas (167.1) and (167.2):

$$(\vartheta)\quad \forall F(Fx \equiv Fy) \rightarrow \Box\forall w(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi^y_x])$$

Now our goal is to show that we can derive the following from our initial assumption:

$$\forall F(Fx \equiv Fy) \rightarrow [\lambda z\, \varphi] = [\lambda z\, \varphi_x^y]$$

So assume $\forall F(Fx \equiv Fy)$ (as a local assumption). It then follows from (ϑ) that $\Box\forall w(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi_x^y])$. Note that by definition (23.2) and the Rule of Biconditional Simplification, the following derivation is valid and modally strict:

$$(\xi)\ \ [\lambda z\, \varphi]{\downarrow}, [\lambda z\, \varphi_x^y]{\downarrow} \vdash_\Box [\lambda z\, \varphi] = [\lambda z\, \varphi_x^y] \equiv \Box\forall w(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi_x^y])$$

But the first premise of this derivation is our initial assumption. And we can derive the second premise from what we know:

> By our initial assumption, $[\lambda z\, \varphi]{\downarrow}$. But given the condition on x in φ, we also know that $[\lambda x\, [\lambda z\, \varphi]{\downarrow}]{\downarrow}$. So we can instantiate $[\lambda x\, [\lambda z\, \varphi]{\downarrow}]$ into our local assumption $\forall F(Fx \equiv Fy)$, to obtain $[\lambda x\, [\lambda z \varphi]{\downarrow}]x \equiv [\lambda x\, [\lambda z \varphi]{\downarrow}]y$. Applying β-Conversion to both sides, this reduces to $[\lambda z\, \varphi]{\downarrow} \equiv [\lambda z\, \varphi]{\downarrow}_x^y$. Hence, $[\lambda z\, \varphi]{\downarrow}_x^y$, i.e., $[\lambda z\, \varphi_x^y]{\downarrow}$.

Since we've established both premises of (ξ), it follows that:

$$[\lambda z\, \varphi] = [\lambda z\, \varphi_x^y] \equiv \Box\forall w(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi_x^y])$$

But since we've established the right-hand side, it follows that $[\lambda z\, \varphi] = [\lambda z\, \varphi_x^y]$. ⋈

(272.5) [**Exercise**] Assume $\forall F(Fx \equiv Fy)$. Now as an instance of axiom (39.2), we know $[\lambda z\, Gx]{\downarrow}$. This implies, by the previous theorem (272.5) with φ set to Gx, that $\forall F(Fx \equiv Fy) \rightarrow [\lambda z\, Gx] = [\lambda z\, Gy]$. Hence, $[\lambda z\, Gx] = [\lambda z\, Gy]$. Now both $Gx{\downarrow}$ and $Gy{\downarrow}$ are also theorems (104.2). So we have established:

$$Gx{\downarrow}\ \&\ Gy{\downarrow}\ \&\ [\lambda z\, Gx] = [\lambda z\, Gy].$$

So by the definition of proposition identity (23.4), $Gx = Gy$. Since G isn't free in our assumption, it follows by GEN that $\forall G(Gx = Gy)$. ($\leftarrow$) By analogous reasoning. ⋈

(273.1) Let φ be the formula $\Box\forall z(z \neq x \rightarrow \exists F\neg(Fz \equiv Fx))$. Then we have to show $[\lambda x\, \varphi]{\downarrow}$. By an appropriate instance of the Kirchner Theorem (271.1), it suffices to show:

$$\Box\forall x\forall y(\forall F(Fx \equiv Fy) \rightarrow (\varphi \equiv \varphi_x^y))$$

By two applications of GEN and then an application of RN, we only need to show, by modally strict means:

$$\forall F(Fx \equiv Fy) \rightarrow (\varphi \equiv \varphi_x^y)$$

So let $\forall F(Fx \equiv Fy)$ be our global assumption, to show both directions of $\varphi \equiv \varphi^y_x$.

($\rightarrow$) Assume φ, i.e.,

(φ) $\Box\forall z(z \neq x \rightarrow \exists F\neg(Fz \equiv Fx))$

We want to show φ^y_x, i.e.,

(φ^y_x) $\Box\forall z(z \neq y \rightarrow \exists F\neg(Fz \equiv Fy))$

Note that φ^y_x can be obtained by the following instance of Rule RN:

$$\text{If } \forall z(z \neq x \rightarrow \exists F\neg(Fz \equiv Fx)) \vdash_\Box \forall z(z \neq y \rightarrow \exists F\neg(Fz \equiv Fy)), \text{ then } \varphi \vdash \varphi^y_x.$$

So if we can show the antecedent of the above instance of RN, then since the premise φ of the consequent is an assumption, we may validly infer the conclusion φ^y_x. To show the antecedent, take the premise of the derivation as an assumption:

(ζ) $\forall z(z \neq x \rightarrow \exists F\neg(Fz \equiv Fx))$

By GEN, we need only derive $z \neq y \rightarrow \exists F\neg(Fz \equiv Fy)$ by modally strict means, since z doesn't occur free in the premise. So assume $z \neq y$. For reductio, assume $\neg\exists F\neg(Fz \equiv Fy)$, i.e., $\forall F(Fz \equiv Fy)$. From this and our global assumption, it follows by biconditional syllogism that $\forall F(Fz \equiv Fx)$. Note independently that it follows from (ζ) that $z \neq x \rightarrow \exists F\neg(Fz \equiv Fx)$, i.e., $z \neq x \rightarrow \neg\forall F(Fz \equiv Fx)$, i.e., $\forall F(Fz \equiv Fx) \rightarrow z = x$. Hence $z = x$. But from this and our local assumption $z \neq y$, it follows that $y \neq x$, on pain of contradiction. And from this last result and (ζ), it follows that $\exists F\neg(Fy \equiv Fx)$. This is equivalent to $\neg\forall F(Fx \equiv Fy)$, which contradicts our global assumption.

($\leftarrow$) By analogous reasoning. ⋈

(273.3) ($\rightarrow$) Assume $D!x$. Then by definition of $D!$ (273.2) and β-Conversion:

$$\Box\forall z(z \neq x \rightarrow \exists F\neg(Fz \equiv Fx))$$

So by the T schema:

$$\forall z(z \neq x \rightarrow \exists F\neg(Fz \equiv Fx))$$

($\leftarrow$) Assume:

(ϑ) $\forall z(z \neq x \rightarrow \exists F\neg(Fz \equiv Fx))$

Since $D!$ exists, it suffices, by β-Conversion and the definition of $D!$, to show:

$$\Box\forall z(z \neq x \rightarrow \exists F\neg(Fz \equiv Fx))$$

By BF, it suffices to show:

$$\forall z \Box (z \neq x \rightarrow \exists F \neg (Fz \equiv Fx))$$

So by GEN, it suffices to show:

$$\Box (z \neq x \rightarrow \exists F \neg (Fz \equiv Fx))$$

And by an instance of (171.2), it suffices to show that both the antecedent and consequent of this necessary conditional are modally collapsed, i.e., it suffices to show both:

(ξ) $\Box (z \neq x \rightarrow \Box z \neq x)$

(ζ) $\Box (\exists F \neg (Fz \equiv Fx) \rightarrow \Box \exists F \neg (Fz \equiv Fx))$

But (ξ) follows by applying RN to (170.2). To obtain (ζ), first apply RN to the modally strict theorem (241) to obtain:

$$\Box (\forall F (Fz \equiv Fx) \rightarrow \Box \forall F (Fz \equiv Fx))$$

But by (171.1), the negation of a modally collapsed formula is also modally collapsed. So:

$$\Box (\neg \forall F (Fz \equiv Fx) \rightarrow \Box \neg \forall F (Fz \equiv Fx))$$

But then (ζ) follows from this by a Rule of Substitution and the modally strict fact that $\neg \forall \alpha \varphi \equiv \exists \alpha \neg \varphi$. ⋈

(273.3) [Alternative proof of ($\leftarrow$) direction, without appealing to (171.1) and (171.2).] ($\leftarrow$) Assume:

(ϑ) $\forall z (z \neq x \rightarrow \exists F \neg (Fz \equiv Fx))$

Since $D!$ exists, it suffices, by β-Conversion and the definition of $D!$, to show:

$$\Box \forall z (z \neq x \rightarrow \exists F \neg (Fz \equiv Fx))$$

By BF, we have to show:

$$\forall z \Box (z \neq x \rightarrow \exists F \neg (Fz \equiv Fx))$$

So by GEN, it suffices to show:

$$\Box (z \neq x \rightarrow \exists F \neg (Fz \equiv Fx))$$

i.e., by a Rule of Substitution and by now familiar reasoning:

$$\Box (\forall F (Fz \equiv Fx) \rightarrow z = x)$$

We can prove this by cases, where the cases are $z = x$ and $z \neq x$.

Case 1: $z = x$. Then $\Box z = x$, by (117.1). So $\Box (\forall F (Fz \equiv Fx) \rightarrow z = x)$, by (158.1).

Case 2: $z \neq x$. Then by (ϑ), $\exists F \neg (Fz \equiv Fx)$. So $\Diamond \exists F \neg (Fz \equiv Fx)$, by T$\Diamond$. Hence, by the 5 schema:

(ξ) $\Box\Diamond\exists F\neg(Fz \equiv Fx)$

Note that $\forall F(Fz \equiv Fx) \to \Box\forall F(Fz \equiv Fx)$ is a modally strict theorem (241). So by contraposition and modal negation, $\Diamond\neg\forall F(Fz \equiv Fx) \to \neg\forall F(Fz \equiv Fx)$. I.e., $\Diamond\exists F\neg(Fz \equiv Fx) \to \exists F\neg(Fz \equiv Fx)$. Since we've derived this result by modally strict means from a modally strict theorem, it follow by RN that:

$$\Box(\Diamond\exists F\neg(Fz \equiv Fx) \to \exists F\neg(Fz \equiv Fx))$$

So by K axiom: $\Box\Diamond\exists F\neg(Fz \equiv Fx) \to \Box\exists F\neg(Fz \equiv Fx)$. So $\Box\exists F\neg(Fz \equiv Fx)$, by ($\xi$). Then, again by (158.1), $\Box(z \neq x \to \exists F\neg(Fz \equiv Fx))$, i.e., $\Box(\forall F(Fz \equiv Fx) \to z = x)$. ⋈

(273.4) Assume $O!x$. By (273.3), it suffices to show $\forall z(z \neq x \to \exists F\neg(Fz \equiv Fx))$, or equivalently, to show $\forall z(\forall F(Fz \equiv Fx) \to z = x)$. So, by GEN, assume $\forall F(Fz \equiv Fx)$. Then by (242.2), $z = x$. ⋈

(273.5) (Exercise)

(273.6) To prove our theorem, we need to find an abstract object that isn't discernible. But theorem (269) ensures that there are at least two of them. For by (269), we know:

$$\exists x\exists y(A!x \,\&\, A!y \,\&\, x \neq y \,\&\, \forall F(Fx \equiv Fy))$$

So let a and b be such objects, so that we know $A!a \,\&\, A!b \,\&\, a \neq b \,\&\, \forall F(Fa \equiv Fb)$. We leave it as an exercise to show that $\neg D!a$ and $\neg D!b$. ⋈

(273.7) Assume $D!x \vee D!y$ and $\forall F(Fx \equiv Fy)$. We then prove $x = y$ by disjunctive syllogism from the two cases of first assumption.

Case 1: $D!x$. Then it follows by (273.3) that $\forall z(z \neq x \to \exists F\neg(Fz \equiv Fx))$, i.e., that $\forall z(\forall F(Fz \equiv Fx) \to z = x)$. Instantiating to y, it follows that $\forall F(Fy \equiv Fx) \to y = x$. But the antecedent of this last result is equivalent to our 2nd assumption. Hence $y = x$, and so by symmetry, $x = y$.

Case 2: $D!y$. By analogous reasoning. ⋈

(273.8) Assume $D!x$. Since the definiens of $D!$ exists, we know by our theory of definition and the definition of $D!$ that:

$$D! = [\lambda x\, \Box\forall z(z \neq x \to \exists F\neg(Fz \equiv Fx))]$$

So by β-Conversion and Rule =E, it is a modally strict theorem that:

(ϑ) $D!x \equiv \Box\forall z(z \neq x \to \exists F\neg(Fz \equiv Fx))$

Hence:

$$\Box\forall z(z \neq x \to \exists F\neg(Fz \equiv Fx)))$$

Then by the 4 schema (165.5):

$$\Box\Box\forall z(z \neq x \rightarrow \exists F \neg(Fz \equiv Fx)))$$

But since (ϑ) is a modally strict theorem, it follows from this by a Rule of Substitution that $\Box D!x$. ⋈

(273.9) – **(273.12)** (Exercises)

(273.13) We want to show: $[\lambda x\, D!x \,\&\, \varphi]\downarrow$. Substituting $D!x \,\&\, \varphi$ for φ in the Kirchner Theorem (271.1), it suffices to show:

$$\Box\forall x\forall y\big(\forall F(Fx \equiv Fy) \rightarrow ((D!x \,\&\, \varphi) \equiv (D!y \,\&\, \varphi^{y}_{x}))\big)$$

By GEN and RN, it suffices to show the embedded conditional. So assume $\forall F(Fx \equiv Fy)$. Without loss of generality, we show only the $\rightarrow$ direction. So assume $D!x \,\&\, \varphi$, to show $D!y \,\&\, \varphi^{y}_{x}$. But $D!y$ follows from $D!x$ and the first assumption. Since $D!x$, we know by (273.3) that $\forall z(z \neq x \rightarrow \exists F \neg(Fz \equiv Fx))$, i.e., $\forall z(\forall F(Fz \equiv Fx) \rightarrow z = x)$. Instantiating to y: $\forall F(Fy \equiv Fx) \rightarrow y = x$. Hence $y = x$, i.e., $x = y$. And so by Rule =E, φ^{y}_{x}. ⋈

(273.14) Assume the antecedent, but for reductio, suppose the consequent is false. Then one or more of the conjuncts in the consequent is false. Suppose the i^{th} conjunct is false, for some i such that $1 \leq i \leq n$. Then we know $\exists G \neg(Gx_i \equiv Gz_i)$. Suppose P is such a property and, without loss of generality, suppose $Px_i \,\&\, \neg Pz_i$ Then consider the n-ary relation $H = [\lambda x_1 \ldots x_i \ldots x_n\, Px_i]$. H clearly exists. To complete our reductio, it remains to show that H is a witness to, and thus implies, $\exists F \neg(Fx_1 \ldots x_n \equiv Fz_1 \ldots z_n)$, for this existential claim contradicts our assumption. So we want to show:

$$\neg(Hx_1 \ldots x_n \equiv Hz_1 \ldots z_n)$$

For reductio, suppose $Hx_1 \ldots x_n \equiv Hz_1 \ldots z_n$. But by the definition and existence of H, and β-Conversion, we know:

$$Hx_1 \ldots x_n \equiv Px_i$$

$$Hz_1 \ldots z_n \equiv Pz_i$$

Hence, $Px_i \equiv Pz_i$, by biconditional reasoning. But this contradicts our assumption that $Px_i \,\&\, \neg Pz_i$. ⋈

(273.15) We want to show, for an arbitrary formula φ, that when $n \geq 1$:

$$[\lambda x_1 \ldots x_n\, D!x_1 \,\&\, \ldots \,\&\, D!x_n \,\&\, \varphi]\downarrow$$

Let φ in the Kirchner Theorem (271.2) be $D!x_1 \,\&\, \ldots \,\&\, D!x_n \,\&\, \varphi$. Then the Kirchner Theorem tells us:

$$\begin{aligned}&[\lambda x_1 \ldots x_n\, D!x_1 \,\&\, \ldots \,\&\, D!x_n \,\&\, \varphi]\downarrow \equiv\\ &\quad \Box\forall x_1 \ldots \forall x_n \forall y_1 \ldots \forall y_n(\forall F(Fx_1 \ldots x_n \equiv Fy_1 \ldots y_n) \rightarrow\\ &\qquad ((D!x_1 \,\&\, \ldots \,\&\, D!x_n \,\&\, \varphi) \equiv (D!y_1 \,\&\, \ldots \,\&\, D!y_n \,\&\, \varphi^{y_1,\ldots,y_n}_{x_1,\ldots,x_n})))\end{aligned}$$

So it suffices to show the right side of the biconditional and, by GEN and RN, it suffices to show:

$$\forall F(Fx_1 \dots x_n \equiv Fy_1 \dots y_n) \rightarrow ((D!x_1 \,\&\, \dots \,\&\, D!x_n \,\&\, \varphi) \equiv (D!y_1 \,\&\, \dots \,\&\, D!y_n \,\&\, \varphi^{y_1,\dots,y_n}_{x_1,\dots,x_n}))$$

So assume:

(ϑ) $\forall F(Fx_1 \dots x_n \equiv Fy_1 \dots y_n)$

Without loss of generality, we show only the ($\rightarrow$) direction. So assume:

(ξ) $D!x_1 \,\&\, \dots \,\&\, D!x_n \,\&\, \varphi$

Note that it follows from (ϑ) by lemma (273.14) that:

(ζ_1) $\forall G(Gx_1 \equiv Gy_1)$

$\vdots$

(ζ_n) $\forall G(Gx_n \equiv Gy_n)$

Thus, we may infer:

$D!y_1$, from the 1st conjunct of (ξ) and (ζ_1)
$\vdots$
$D!y_n$, from the n^{th} conjunct of (ξ) and (ζ_n)

So it remains to show $\varphi^{y_1,\dots,y_n}_{x_1,\dots,x_n}$. But we may also infer:

$x_1 = y_1$, from the 1st conjunct of (ξ), (ζ_1), and (273.7)
$\vdots$
$x_n = y_n$, from the n^{th} conjunct of (ξ), (ζ_n), and (273.7)

But from these results and the final conjunct of (ξ) it follows that $\varphi^{y_1,\dots,y_n}_{x_1,\dots,x_n}$ by Rule =E. ⋈

(273.16) This is an instance of (273.15). ⋈

(273.18) Since $[\lambda xy\, D!x \,\&\, D!y \,\&\, x = y]\!\downarrow$ (273.16), we know both (ξ), which follows from the theory of definitions-by-= and the definition of $=_D$ (273.17), and (ϑ), which follows by β-Conversion:

(ξ) $=_D \,=\, [\lambda xy\, D!x \,\&\, D!y \,\&\, x = y]$

(ϑ) $[\lambda xy\, D!x \,\&\, D!y \,\&\, x = y]xy \equiv D!x \,\&\, D!y \,\&\, x = y$

Hence, by Rule =E and the application of infix notation for $=_D$:

$$x =_D y \equiv D!x \,\&\, D!y \,\&\, x = y \qquad ⋈$$

(273.19) (Exercise)

(273.20) ($\rightarrow$) Assume $x =_D y$. Then by (273.18), $D!x \,\&\, D!y \,\&\, x = y$. So it remains to show $\Box\forall F(Fx \equiv Fy)$. But this follows from the easy theorem $\Box\forall F(Fx \equiv Fx)$ and $x = y$. ($\leftarrow$) Assume $D!x \,\&\, D!y \,\&\, \Box\forall F(Fx \equiv Fy)$. Then by the right-to-left direction of (273.18), it remains only to show $x = y$. But this follows from the first and third conjuncts of our assumption by (273.7). ⋈

(273.21) By the T schema, it suffices to show only the ($\rightarrow$) direction. Assume $x =_D y$. Then by (273.18):

$$D!x \,\&\, D!y \,\&\, x = y$$

But by (273.8), the first two conjuncts imply, respectively, $\Box D!x$ and $\Box D!y$. And by (125.1), the third conjunct implies $\Box x = y$. Hence:

$$\Box D!x \,\&\, \Box D!y \,\&\, \Box x = y$$

So by (158.3):

$$\Box(D!x \,\&\, D!y \,\&\, x = y)$$

Since (273.18) is a modally strict theorem, it follows by a Rule of Substitution that $\Box x =_D y$. ⋈

(273.22) – **(273.25)** (Exercises)

(273.26) By applying RN to the left-to-right direction of (273.21) and applying an appropriate instance of (171.1). ⋈

(273.27) – **(273.29)** (Exercises)

(273.30) Assume $D!x$. Then since $x = x$ is a modally strict theorem (117.1), we know:

$$D!x \,\&\, D!x \,\&\, x = x$$

Hence by (273.18), $x =_D x$. ⋈

(273.31) Assume $x =_D y$. Then by (273.18):

$$D!x \,\&\, D!y \,\&\, x = y$$

We leave it as an exercise to show that this implies, by modally strict reasoning:

$$D!y \,\&\, D!x \,\&\, y = x$$

Hence, $y =_D x$, by (273.18). ⋈

(273.32) Assume $x =_D y$ and $y =_D z$. Then by (273.18), we know both:

$$D!x \,\&\, D!y \,\&\, x = y$$

$D!y \,\&\, D!z \,\&\, y = z$

We leave it as an exercise to show that these imply, by modally strict reasoning:

$D!x \,\&\, D!z \,\&\, x = z$

Hence, $x =_D z$, by (273.18). ⋈

(273.33) We first independently establish, as a modally strict theorem, that:

$D!x \to (x = y \equiv x =_D y)$

Proof. Assume $D!x$. ($\to$) Assume $x = y$. Since $D!x$, it follows from (273.30) that $x =_D x$. Hence, by Rule =E, $x =_D y$. ($\leftarrow$) Assume $x =_D y$. Then by (273.18), $x = y$.

Given this modally strict proof, it follows by Rule RM that:

(ϑ) $\Box D!x \to \Box(x = y \equiv x =_D y)$

is a theorem. Now to prove our theorem, we reason by cases from $D!x \vee D!y$.

Case 1. Assume $D!x$. Then $\Box D!x$, by (273.8). Hence, $\Box(x = y \equiv x =_D y)$ by (ϑ).

Case 2. By analogous reasoning. ⋈

(273.34) [The following reasoning is analogous to that used in (240.2).] Assume $D!y$. Note independently that our conclusion is the consequent of the following instance of axiom (49):

$([\lambda x\, x =_D y]{\downarrow} \,\&\, \Box\forall x(x =_D y \equiv x = y)) \to [\lambda x\, x = y]{\downarrow}$

So it suffices to show:

(ϑ) $[\lambda x\, x =_D y]{\downarrow}$

(ξ) $\Box\forall x(x =_D y \equiv x = y)$

But (ϑ) is axiomatic (39.2) since $[\lambda x\, x =_D y]$ is a core λ-expression. To show (ξ), note that our assumption $D!y$ implies $D!x \vee D!y$, and so it follows from (273.33) that $\Box(x = y \equiv x =_D y)$. Since it is a modally strict theorem that biconditionals commute, it follows by a Rule of Substitution that $\Box(x =_D y \equiv x = y)$. Since x isn't free in our assumption, it follows by GEN that $\forall x \Box(x =_D y \equiv x = y)$. So by the Barcan Formula (167.1), it follows that $\Box\forall x(x = y \equiv x =_D y)$. ⋈

(273.35) Assume $D!x$, $D!y$. Note that both $[\lambda z\, z =_D x]{\downarrow}$ and $[\lambda z\, z =_D y]{\downarrow}$, by now familiar reasoning. ($\to$) Assume $x \neq y$. Then $x \neq_D y$, by (273.19). For reductio, assume $[\lambda z\, z =_D x] = [\lambda z\, z =_D y]$. Since $D!x$, we know by the reflexivity of $=_D$ (273.30) that $x =_D x$. Since $[\lambda z\, z =_D x]{\downarrow}$ and $x{\downarrow}$, it follows by Rule $\overleftarrow{\beta}$C (184.2.a) that $[\lambda z\, z =_D x]x$. But then by Rule =E, it follows that $[\lambda z\, z =_D y]x$. By Rule $\overrightarrow{\beta}$C

(184.1.a), it follows that $x =_D y$. But by (273.25), this contradicts $x \neq_D y$. ($\leftarrow$) The proof of this direction, which is equivalent to its contrapositive $x = y \rightarrow [\lambda z\, z =_D x] = [\lambda z\, z =_D y]$, is trivial. ⋈

(276.1) $[\lambda y\, p]$ is a core λ-expression and so $[\lambda y\, p]\downarrow$ is an instance of axiom (39.2). So the present theorem is *axiomatic*, since it is a universal closure of this instance. ⋈

(276.2) Note that this schema has *two* metavariables, ν and φ. As mentioned in the text, there are two interesting ways to prove this theorem. First, it can be derived from the previous theorem, (276.1), by the following reasoning. Since ν is a *metavariable* ranging over individual variables, note that the following is a *schema* and that its instances are alphabetic variants of (276.1) and so theorems:

(ϑ) $\forall p([\lambda \nu\, p]\downarrow)$

Now, independently, we know that $\varphi\downarrow$ for any φ (104.2). *A fortiori*, $\varphi\downarrow$ for any φ in which ν doesn't occur free. So a formula φ in which ν doesn't occur free is substitutable for the variable p in the matrix of (ϑ). Hence, by $\forall$E, $[\lambda \nu\, \varphi]\downarrow$, for any formula φ in which ν doesn't occur free.

The following, rather different, proof is simpler (despite being longer) and may come as a surprise if one has forgotten the definition of $\downarrow$ as it applies to propositions. By our conventions for definitions (17.2), definition (20.3) is an easier-to-read version of the following definition, in which Π^0 is a metavariable ranging over 0-ary relation terms:

$$\Pi^0\downarrow \equiv_{df} [\lambda \nu\, \Pi^0]\downarrow, \text{ provided } \nu \text{ doesn't occur free in } \Pi^0$$

So by Rule $\equiv_{df}$E, we know the following are theorems:

$$\Pi^0\downarrow \equiv [\lambda \nu\, \Pi^0]\downarrow, \text{ provided } \nu \text{ doesn't occur free in } \Pi^0)$$

But since all and only formulas are 0-ary relation terms, we may rewrite the above as:

(ξ) $\varphi\downarrow \equiv [\lambda \nu\, \varphi]\downarrow$, provided ν doesn't occur free in φ)

So consider any formula φ in which the individual variable ν doesn't occur free. Then $\varphi\downarrow$ by (104.2). So by (ξ), it follows that $[\lambda \nu\, \varphi]\downarrow$. ⋈

(276.3) From the fact that $[\lambda y\, p]\downarrow$ (39.2), Strengthened β-Conversion (181) implies $[\lambda y\, p]x \equiv p$. So by GEN, it follows that $\forall x([\lambda y\, p]x \equiv p)$. Since this last theorem is modally strict, it follows by RN that $\Box\forall x([\lambda y\, p]x \equiv p)$. Now assume $F = [\lambda y\, p]$. Then by substitution of identicals into what we've already established, it follows that $\Box\forall x(Fx \equiv p)$. ⋈

(276.4) Assume *Propositional*(F). Then by definition, $\exists p(F = [\lambda y\, p])$. Assume q_1 is such a proposition, so that we know $F = [\lambda y\, q_1]$. Hence by (125.2), $\Box(F =$

$[\lambda y\, q_1])$. So, by $\exists$I and $\exists$E, $\exists p \Box(F = [\lambda y\, p])$. Then by the Buridan Formula (168.1), $\Box \exists p(F = [\lambda y\, p])$. Hence, $\Box$*Propositional*(F), by definition (275) and the Rule of Substitution for Defined Formulas (160.3). ⋈

(278) Assume *Propositional*(F). Since $F\!\downarrow$, it remains only to show $\Box(\exists x Fx \to \forall x Fx)$, by definition (277) of *Indiscriminate*(F) and Rule $\equiv_{df}$I. By (276.4), our assumption implies that $\Box$*Propositional*(F). So by definition (275) and the Rule of Substitution for Defined Formulas (160.3), it follows that $\Box \exists p(F = [\lambda y\, p])$. Now if we can establish:

$$\Box \exists p(F = [\lambda y\, p]) \to \Box(\exists x Fx \to \forall x Fx)$$

we're done. But by Rule RM (157.1), it suffices to give a modally strict proof of $\exists p(F = [\lambda y\, p]) \to (\exists x Fx \to \forall x Fx)$. So assume $\exists p(F = [\lambda y\, p])$. Suppose q_1 is such a proposition, so that we know $F = [\lambda y\, q_1]$. Now assume $\exists x Fx$, to show $\forall x Fx$. Suppose a is such an object, so that we know Fa. Then $[\lambda y\, q_1]a$, and by Rule $\overrightarrow{\beta}$C (184.1.a), q_1. But since $[\lambda y\, q_1]\!\downarrow$ and $x\!\downarrow$, it follows that $[\lambda y\, q_1]x$, by Rule $\overleftarrow{\beta}$C (184.2.a). So Fx. Since x doesn't occur free in any of our assumptions, we may infer $\forall x Fx$ by GEN. ⋈

(279.1) As an instance of (38.1), we know:

$$\forall x Fx \to (\exists x Fx \to \forall x Fx)$$

Since this theorem is modally strict, it follows by RM (157.1) that:

$$\Box \forall x Fx \to \Box(\exists x Fx \to \forall x Fx)$$

By definition (200.1) and the Rule of Substitution for Defined Formulas (160.3), we may infer:

$$(\vartheta)\quad \textit{Necessary}(F) \to \Box(\exists x Fx \to \forall x Fx)$$

Independently, it follows from definition (277), the axiom $F\!\downarrow$, and Rule $\equiv$S of Biconditional Simplification, that the following is a modally strict theorem:

$$\textit{Indiscriminate}(F) \equiv \Box(\exists x Fx \to \forall x Fx)$$

From this and (ϑ), it follows that that *Necessary*$(F) \to$ *Indiscriminate*(F), by the Rule of Substitution (160.2). ⋈

(279.2) As an instance of (77.3), we know:

$$\neg \exists x Fx \to (\exists x Fx \to \forall x Fx)$$

So from the right-to-left direction of an appropriate instance of (103.4) and the above, it follows by hypothetical syllogism that:

$$\forall x \neg Fx \to (\exists x Fx \to \forall x Fx)$$

Since this theorem is modally strict, it follows by RM (157.1) that:

$\Box\forall x\neg Fx \to \Box(\exists xFx \to \forall xFx)$

Independently, we know that $Impossible(F) \equiv \Box\forall x\neg Fx$ is a modally strict theorem, by definition (200.2), the axiom $F\downarrow$ and Rule $\equiv$S (91.1). So our last two results imply the following, by the Rule of Substitution (160.2):

$Impossible(F) \to \Box(\exists xFx \to \forall xFx)$

Again, independently, we know $Indiscriminate(F) \equiv \Box(\exists xFx \to \forall xFx)$, by now familiar reasoning. So by the Rule of Substitution, our last two results imply:

$Impossible(F) \to Indiscriminate(F)$ ⋈

(279.3.a) Suppose, for reductio, that $E!$ is indiscriminate. Then, by definition (277) and Rule $\equiv_{df}$E:

(ϑ) $E!\downarrow \,\&\, \Box(\exists xE!x \to \forall xE!x)$

Now independently we know $\Diamond\exists xE!x$ (205.3). But this and the right conjunct of (ϑ) imply $\Diamond\forall xE!x$, by (158.13). Moreover, we also know independently that $\Box\neg\forall xE!x$ (227.5), which is equivalent to $\neg\Diamond\forall xE!x$. Contradiction. ⋈

(279.3.b) Suppose, for reductio, that $\overline{E!}$ is indiscriminate. Then, by definition (277) and Rule $\equiv_{df}$E:

(ϑ) $\overline{E!}\downarrow \,\&\, \Box(\exists x\overline{E!}x \to \forall x\overline{E!}x)$

Now independently, we know $\Box\neg\forall xE!x$ (227.5), which is equivalent to $\Box\exists x\neg E!x$ (exercise). In turn, this is equivalent to $\Box\exists x\overline{E!}x$ (exercise). But from this last result and the right conjunct of (ϑ), it follows that $\Box\forall x\overline{E!}x$, by the K schema. Independently, however, we know $\Diamond\exists xE!x$ (205.3), i.e., by definition, $\neg\Box\neg\exists xE!x$. But this is equivalent to $\neg\Box\forall x\neg E!x$ (exercise), which in turn is equivalent to $\neg\Box\forall x\overline{E!}x$ (exercise). Contradiction. ⋈

(279.3.c) Suppose, for reductio, that $O!$ is indiscriminate. Then, by definition (277) and Rule $\equiv_{df}$E:

(ϑ) $O!\downarrow \,\&\, \Box(\exists xO!x \to \forall xO!x)$

Now independently we know $\Box\exists xO!x$ (227.1). From this and the right conjunct of (ϑ), it follows by the K schema that $\Box\forall xO!x$. By the T schema, $\forall xO!x$. But an application of the T schema to (227.3) yields $\neg\forall xO!x$. Contradiction. ⋈

(279.3.d) Suppose, for reductio, that $A!$ is indiscriminate. Then, by definition (277) and Rule $\equiv_{df}$E:

(ϑ) $A!\downarrow \,\&\, \Box(\exists xA!x \to \forall xA!x)$

Now independently we know $\Box\exists x A!x$ (227.1). From this and the right conjunct of (ϑ), it follows by the K schema that $\Box\forall x A!x$. By the T schema, $\forall x A!x$. But an application of the T schema to (227.4) yields $\neg\forall x A!x$. Contradiction. ⋈

(279.4.a) – **(279.4.d)** (Exercises)

(280.1) Assume $\Diamond\exists p(F = [\lambda y\, p])$, to show $\exists p(F = [\lambda y\, p])$. By BF$\Diamond$ (167.3), it follows that $\exists p\Diamond(F = [\lambda y\, p])$. Assume q_1 is an arbitrary such p, so that we know $\Diamond(F = [\lambda y\, q_1])$. But properties that are possibly identical are identical (170.1), and since $F\downarrow$ and $[\lambda y\, q_1]\downarrow$, it follows that $F = [\lambda y\, q_1]$. Hence, by $\exists$I (101.2) that $\exists p(F = [\lambda y\, p])$. ⋈

(280.2) Assume $\forall p(F \neq [\lambda y\, p])$. Then $F \neq [\lambda y\, p]$. Since $F\downarrow$ and $[\lambda y\, p]\downarrow$, it follows by (170.2) that $\Box F \neq [\lambda y\, p]$. Hence, by GEN, $\forall p \Box F \neq [\lambda y\, p]$. And by BF (167.1), $\Box\forall p(F \neq [\lambda y\, p])$. ⋈

(280.3) From (280.1) by (166.1), given that the (280.1) is a modally-strict theorem. ⋈

(280.4) From (280.2) by (166.2), given that (280.2) is a modally strict theorem. ⋈

(281.1) Assume $\Diamond\forall F(xF \to \exists p(F = [\lambda y\, p]))$, for conditional proof. By the Buridan$\Diamond$ formula (168.2), this implies:

(ϑ) $\forall F\Diamond(xF \to \exists p(F = [\lambda y\, p]))$

Now we want to show $\forall F(xF \to \exists p(F = [\lambda y\, p]))$. By GEN it suffices to show $xF \to \exists p(F = [\lambda y\, p])$. So assume xF. Then $\Box xF$, by axiom (51). Now if we instantiate (ϑ) to F, it follows that $\Diamond(xF \to \exists p(F = [\lambda y\, p]))$. This together with $\Box xF$ yields $\Diamond\exists p(F = [\lambda y\, p])$, by (162.4). But by a previous theorem (280.1), it follows that $\exists p(F = [\lambda y\, p])$. ⋈

(281.2) From (281.1), by (166.1), given that (281.1) is a $\Box$-theorem. ⋈

(288.1) As an instance of the Comprehension Principle for Abstract Objects (53), we know:

$$\exists x(A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])))$$

But by definition (286) and the Rule of Substitution for Defined Formulas (160.3), it follows that:

$$\exists x TruthValueOf(x, p) \qquad ⋈$$

(288.2) By the same reasoning as in (288.1), but starting with an instance of the Strengthened Comprehension for Abstract Objects (250). ⋈

(288.3) Assume:

(ϑ) $\mathit{TruthValueOf}(x, p)$

(ξ) $\mathit{TruthValueOf}(y, q)$

($\rightarrow$) Assume $x = y$. By (ϑ) and definition (286), we know:

$$A!x \,\&\, \forall F(xF \equiv \exists r((r \equiv p) \,\&\, F = [\lambda z\, r]))$$

Since it is easy to establish independently that $(p \equiv p) \& [\lambda z p] = [\lambda z p]$, it follows that $\exists r((r \equiv p) \,\&\, [\lambda z\, p] = [\lambda z\, r])$. So $x[\lambda y\, p]$. From this and our local assumption $x = y$, it follows that $y[\lambda z\, p]$. Since (ξ) similarly implies $\forall F(yF \equiv \exists r((r \equiv q) \,\&\, F = [\lambda z\, r]))$, it follows that $\exists r((r \equiv q) \,\&\, [\lambda z\, p] = [\lambda z\, r])$. Suppose r_1 is such a proposition, so that we know $(r_1 \equiv q) \,\&\, [\lambda z\, p] = [\lambda z\, r_1]$. The second conjunct implies, by the identity conditions for propositions (23.4), that $p = r_1$. So $p \equiv q$.

($\leftarrow$) Assume $p \equiv q$. Since x and y are both abstract, it suffices to show $\forall G(xG \equiv yG)$. So, by GEN, we show $xG \equiv yG$. ($\rightarrow$) Assume xG. Then by (ϑ), and definition (286), we know: $\exists r((r \equiv p) \,\&\, G = [\lambda z\, r])$. Suppose $(r_2 \equiv p) \,\&\, G = [\lambda z\, r_2]$. Then $(r_2 \equiv q) \,\&\, G = [\lambda z\, r_2]$. So $\exists r((r \equiv q) \,\&\, G = [\lambda z\, r])$ and, hence, yG, by (ξ) and definition (286). ($\leftarrow$) By analogous reasoning. ⋈

(289.1) ($\rightarrow$) Assume p. By GEN, it suffices to show the equivalence of (ϑ) and (ξ):

(ϑ) $\exists q(q \,\&\, F = [\lambda y\, q])$

(ξ) $\exists q((q \equiv p) \,\&\, F = [\lambda y\, q])$

We show each direction separately:

($\rightarrow$) Assume (ϑ). Suppose q_1 is such a proposition, so that we know q_1 is true and $F = [\lambda y\, q_1]$. But since p is true, it is materially equivalent to q_1. So it follows that $(q_1 \equiv p) \,\&\, F = [\lambda y\, q_1]$, and by $\exists$I, that (ξ).

($\leftarrow$) Assume (ξ). Suppose q_2 is an arbitrary such proposition, so that we know $q_2 \equiv p$ and $F = [\lambda y q_2]$. But then since p is true, q_2 is true. So $q_2 \& F = [\lambda y q_2]$, and by $\exists$I, (ϑ).

($\leftarrow$) Our assumption is:

$$\forall F[\exists q(q \,\&\, F = [\lambda y\, q]) \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])]$$

We want to show p. For reductio, assume $\neg p$. Since $[\lambda y\, \neg p]\downarrow$, instantiate the property $[\lambda y\, \neg p]$ into our assumption, to obtain:

(ζ) $\exists q(q \,\&\, [\lambda y\, \neg p] = [\lambda y\, q]) \equiv \exists q((q \equiv p) \,\&\, [\lambda y\, \neg p] = [\lambda y\, q])$

Note also that Rule =I (118.2) governs $[\lambda y\, \neg p]$, and so we know $[\lambda y\, \neg p] = [\lambda y\, \neg p]$. Hence we know:

$$\neg p \,\&\, [\lambda y\, \neg p] = [\lambda y\, \neg p]$$

So, by ∃I,

$$\exists q(q \,\&\, [\lambda y\, \neg p] = [\lambda y\, q])$$

From this last conclusion and (ζ), it follows that:

$$\exists q((q \equiv p) \,\&\, [\lambda y\, \neg p] = [\lambda y\, q])$$

Let q_1 be such a proposition, so that we know $(q_1 \equiv p) \,\&\, [\lambda y\, \neg p] = [\lambda y\, q_1]$. By the definition of identity for propositions, the second conjunct implies $\neg p = q_1$. But $\neg p$ is true by our reductio assumption. So it follows that q_1 is true. But since q_1 is equivalent to p, p is true. Contradiction. ⋈

(289.2) (→) Assume $\neg p$. By GEN, it suffices to show the equivalence of (ϑ) and (ξ):

(ϑ) $\exists q(\neg q \,\&\, F = [\lambda y\, q])$

(ξ) $\exists q((q \equiv p) \,\&\, F = [\lambda y\, q])$

We show each direction separately:

> (→) Assume (ϑ). Suppose q_1 is an arbitrary such proposition, so that we know $\neg q_1$ and $F = [\lambda y\, q_1]$. But since $\neg p$, it follows that $q_1 \equiv p$. Hence $(q_1 \equiv p) \,\&\, F = [\lambda y\, q_1]$, and by ∃I, that ($\xi$).
>
> (←) Assume (ξ). Suppose q_2 is an arbitrary such proposition, so that we know $q_2 \equiv p$ and $F = [\lambda y\, q_2]$. But then since $\neg p$, it follows that $\neg q_2$. So $\neg q_2 \,\&\, F = [\lambda y\, q_2]$, and by ∃I, ($\vartheta$).

(←) (Exercise) ⋈

(291.1) Assume:

(ϑ) $A!x \,\&\, \forall F(xF \equiv \exists q(q \,\&\, F = [\lambda y\, q]))$

To show $TruthValue(x)$, we have to show, by definition (290):

$$\exists p(TruthValueOf(x, p))$$

We can do this if we take our witness to be $\forall x(E!x \to E!x)$. Call this proposition p_0 (we also defined p_0 this way in (208). So we have to show:

$$TruthValueOf(x, p_0)$$

By definition (286), we have to show:

(ξ) $A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p_0) \,\&\, F = [\lambda y\, q]))$

Now clearly it follows from the tautology $E!x \rightarrow E!x$, by GEN, that p_0 is provably true, and by modally strict means. So it follows from (289.1) that:

$$\forall F[\exists q(q \,\&\, F = [\lambda y\, q]) \equiv \exists q((q \equiv p_0) \,\&\, F = [\lambda y\, q])]$$

Then by ∀E:

$$\exists q(q \,\&\, F = [\lambda y\, q]) \equiv \exists q((q \equiv p_0) \,\&\, F = [\lambda y\, q])$$

Since we've established this by a modally-strict proof, it follows from (ϑ) by the Rule of Substitution (160.2) that (ξ). ⋈

(291.2) (Exercise)

(292) Consider the following two instances of Object Comprehension (53):

$$\exists x(A!x \,\&\, \forall F(xF \equiv \exists p(p \,\&\, F = [\lambda y\, p])))$$

$$\exists x(A!x \,\&\, \forall F(xF \equiv \exists p(\neg p \,\&\, F = [\lambda y\, p]))$$

Let a, b be arbitrary such objects, respectively, so that we know:

(ϑ) $A!a \,\&\, \forall F(aF \equiv \exists p(p \,\&\, F = [\lambda y\, p]))$

(ξ) $A!b \,\&\, \forall F(bF \equiv \exists p(\neg p \,\&\, F = [\lambda y\, p]))$

We now develop modally-strict arguments that show:

(1) $\mathit{TruthValue}(a) \,\&\, \mathit{TruthValue}(b)$

(2) $a \neq b$

(3) $\forall z(\mathit{TruthValue}(z) \rightarrow z = a \vee z = b)$

Note that (1) follows from (291.1) and (291.2), respectively, given (ϑ) and (ξ).

To show (2), note that the left conjuncts of (ϑ) and (ξ), respectively, are that $A!a$ and $A!b$. So, it suffices, by theorem (245.2), to show a encodes a property that b fails to encode, or vice versa. Consider the property $[\lambda y\, p_0]$, where p_0 is the proposition $\forall x(E!x \rightarrow E!x)$. Now although p_0 is a defined constant, we can cite (39.2) to conclude $[\lambda y\, p_0]\downarrow$, since $[\lambda y\, p_0]$ still qualifies as a core λ-expression. And since $[\lambda y\, p_0]\downarrow$, it follows from the right conjunct of (ϑ) that:

$$a[\lambda y\, p_0] \equiv \exists p(p \,\&\, [\lambda y\, p_0] = [\lambda y\, p])$$

But the right condition is easily derived since p_0 is a witness: it is true and $[\lambda y\, p_0] = [\lambda y\, p_0]$ is known by the special case of Rule =I (118.2). So $a[\lambda y\, p_0]$. Now, for reductio, assume $b[\lambda y\, p_0]$. Then by the right conjunct of (ξ), it follows that $\exists p(\neg p \,\&\, [\lambda y\, p_0] = [\lambda y\, p])$. Suppose p_1 is an arbitrary such proposition, so that we know $\neg p_1 \,\&\, [\lambda y\, p_0] = [\lambda y\, p_1]$. By the theorem governing proposition identity (116.3), it follows that $p_0 = p_1$. Hence by Rule =E (110), it follows that $\neg p_0$, which contradicts the fact that p_0 is true.

To show (3), since z isn't free in any assumption, it suffices by GEN to show $TruthValue(z) \to z = a \vee z = b$. Assume $TruthValue(z)$. Then $\exists p(TruthValueOf(z, p))$. Let p_2 be such a proposition, so that we know $TruthValueOf(z, p_2)$. Then by (286), we know:

(ζ) $A!z \,\&\, \forall F(zF \equiv \exists q((q \equiv p_2) \,\&\, F = [\lambda y\, q]))$

Now we reason by disjunctive syllogism, from the tautology $p_2 \vee \neg p_2$, to show $z = a \vee z = b$.

- Assume p_2, to show $z = a$. Since both z and a are abstract, it suffices to show $\forall G(zG \equiv aG)$ by theorem (245.2). Since G isn't free in our assumption, it suffices, by GEN, to show $zG \equiv aG$. But since p_2 is true, we may reason as follows:

$$\begin{array}{llll} zG & \equiv & \exists q((q \equiv p_2) \,\&\, G = [\lambda y\, q]) & \text{by right conjunct of } (\zeta) \\ & \equiv & \exists q(q \,\&\, G = [\lambda y\, q]) & \text{by (289.1) and the truth of } p_2 \\ & \equiv & aG & \text{by right conjunct of } (\vartheta) \end{array}$$

- Assume $\neg p_2$ to show $z = b$ by analogous reasoning, using (289.2).

Hence, by disjunctive syllogism, $z = a \vee z = b$.

Since the proofs of (1), (2), and (3) make no appeal to any $\star$-theorems, they constitute a modally-strict proof that there are exactly two truth-values. ⋈

(293) It follows from (288.2) that $\mathcal{A}\exists! x TruthValueOf(x, p)$, by the Rule of Actualization. Hence by (176.2), $\iota x TruthValueOf(x, p)\downarrow$. ⋈

(296.1) By (293), we know:

(ϑ) $\iota x TruthValueOf(x, p)\downarrow$

By Rule $\equiv$Df, definition (286) yields the following as a theorem:

$$TruthValueOf(x, p) \equiv A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q]))$$

So by GEN and RN:

(ξ) $\Box\forall x(TruthValueOf(x, p) \equiv A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])))$

It follows from (ϑ) and (ξ) by (149.3) that:

(ζ) $\iota x TruthValueOf(x, p) = \iota x(A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])))$

Since all of the terms in question are significant, it follows by (294), (ζ), and the transitivity of identity that:

$$\circ p = \iota x(A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])))$$ ⋈

(296.2) Since $p \equiv p$ is a tautology (88.3.a) and $[\lambda y\, p] = [\lambda y\, p]$ is a modally strict theorem (118.2), we can conjoin them and existentially generalize to conclude:

$$\exists q((q \equiv p) \,\&\, [\lambda y\, p] = [\lambda y\, q])$$

From this theorem, we may conclude, by the Rule of Actualization:

$$(\vartheta)\quad \mathcal{A}\exists q((q \equiv p) \,\&\, [\lambda y\, p] = [\lambda y\, q])$$

Now if we let φ be the formula $\exists q((q \equiv p) \,\&\, F = [\lambda y\, q])$, then (ϑ) has the form:

$$\mathcal{A}\varphi_F^{[\lambda y\, p]}$$

But by (258.2), we know:

$$\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))[\lambda y\, p] \equiv \mathcal{A}\varphi_F^{[\lambda y\, p]}$$

Hence $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))[\lambda y\, p]$. But since φ is $\exists q((q \equiv p) \,\&\, [\lambda y\, p] = [\lambda y\, q])$, it follows by (296.1) and Rule =E that $\circ p[\lambda y\, p]$. So by definition (295), $\circ p \Sigma p$. ⋈

(297) We can establish our theorem by taking as our witnesses a contingently true proposition and a propositional property constructed out of a necessarily true proposition. We know that there are contingently true propositions by (217.1). So assume p_1 is such a proposition, so that we know $p_1 \,\&\, \Diamond\neg p_1$.[458] Furthermore, where p_0 is the necessarily true proposition $\forall x(E!x \to E!x)$ (208.1), consider the property $[\lambda y\, p_0]$. Clearly, then, $[\lambda y\, p_0]\downarrow$. By two applications of ∃Iand an application of ∃E, it suffices to show:

$$\Diamond(\exists q((q \equiv p_1) \,\&\, [\lambda y\, p_0] = [\lambda y\, q]) \,\&\, \Diamond\neg\exists q((q \equiv p_1) \,\&\, [\lambda y\, p_0] = [\lambda y\, q]))$$

By the T◊ schema, it suffices to show:

$$(\zeta)\quad \exists q((q \equiv p_1) \,\&\, [\lambda y\, p_0] = [\lambda y\, q]) \,\&\, \Diamond\neg\exists q((q \equiv p_1) \,\&\, [\lambda y\, p_0] = [\lambda y\, q])$$

The first conjunct of (ζ) is easy to establish, since p_0 is a witness: p_0 is materially equivalent to p_1 (given that p_0 and p_1 are both true) and $[\lambda y\, p_0] = [\lambda y\, p_0]$ follows from $[\lambda y\, p_0]\downarrow$ by Rule =I (118.1) (Variant).

To establish the second conjunct of (ζ), we first prove the following general lemma, for any propositions p and r:

$$(\text{A})\quad \Box[\exists q((q \equiv p) \,\&\, [\lambda y\, r] = [\lambda y\, q]) \to (r \equiv p)]$$

[458] We've noted on prior occasions why reasoning of the kind used in the remainder of this proof is modally strict notwithstanding the appeal to assumptions that aren't necessary truths. In this particular case, the assumption will be discharged by ∃E – we prove our theorem follows from the assumption $p_1 \,\&\, \Diamond\neg p_1$ and then conclude by ∃E that our theorem follows from the modally strict theorem $\exists p(p \,\&\, \Diamond\neg p)$. See the discussion in Remark (70), Remark (218), and the proof of theorem (221.1).

Proof. By RN it suffices to give a modally strict proof of: $\exists q((q \equiv p) \,\&\, [\lambda y\, r] = [\lambda y\, q]) \to (p \equiv r)$. So assume $\exists q((q \equiv p) \,\&\, [\lambda y\, r] = [\lambda y\, q])$. Suppose s is such a proposition, so that we know both $s \equiv p$ and $[\lambda y\, r] = [\lambda y\, s]$. The latter implies $r = s$. Hence, $r \equiv p$.

As an instance of (A), we know:

(B) $\Box[\exists q((q \equiv p_1) \,\&\, [\lambda y\, p_0] = [\lambda y\, q]) \to (p_0 \equiv p_1)]$

Now to prove the second conjunct of (ζ), assume its negation for reductio, i.e., assume $\Box\exists q((q \equiv p_1) \,\&\, [\lambda y\, p_0] = [\lambda y\, q])$. Then from this and (B), it follows by the K axiom (45.1) that $\Box(p_0 \equiv p_1)$. But by (158.6), this result and our assumption $\Box p_0$ imply $\Box p_1$, which contradicts the assumption that $\Diamond\neg p_1$.[459] ⋈

(299.1)★ By theorem (293) and definition (294), the Rule of Identity by Definition (120) implies:

(ϑ) $\circ p = \imath x TruthValueOf(x, p)$

This implies $\circ p\!\downarrow$, by (107.1). Hence, by (145.2)★:

$$TruthValueOf(\circ p, p) \qquad ⋈$$

(299.2)★ By the previous theorem (299.1)★ and definition (286), it follows that $A!\circ p \,\&\, \forall F(\circ pF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q]))$. Our theorem follows *a fortiori*. ⋈

(299.3)★ By instantiating (299.2)★ to the property $[\lambda y\, r]$, it follows that:

(ϑ) $\circ p[\lambda y\, r] \equiv \exists q((q \equiv p) \,\&\, [\lambda y\, r] = [\lambda y\, q])$

Now we have to show $\circ p\Sigma r \equiv (r \equiv p)$. ($\to$) Assume $\circ p\Sigma r$. By definition (295), it follows that $\circ p[\lambda y\, r]$. From this and (ϑ), we may conclude that:

(ξ) $\exists q((q \equiv p) \,\&\, [\lambda y\, r] = [\lambda y\, q])$

Let q_1 be an arbitrary such proposition, so that we know:

(ζ) $(q_1 \equiv p) \,\&\, [\lambda y\, r] = [\lambda y\, q_1]$

From the right conjunct of (ζ), it follows by the theorem governing proposition identity (116.3) that $r = q_1$, which by symmetry yields $q_1 = r$. But from this and the left conjunct of (ζ), it follows by Rule =E that $r \equiv p$.

[459] Intuitively, this reductio proof can be restated more simply as follows. For reductio, assume the negation of the second conjunct of (ζ), i.e., assume that necessarily, some proposition materially equivalent to p_1, say r, is such that $[\lambda y\, p_0] = [\lambda y\, r]$. By definition of proposition identity, this is just to suppose that necessarily, r is both materially equivalent to p_1 and identical to p_0. But now we have a contradiction: if *necessarily* r is both materially equivalent to p_1 and identical to a necessary truth, then p_1 is a necessary truth, contradicting the hypothesis that p_1 is contingently true.

($\leftarrow$) Suppose that $r \equiv p$. Independently, since $[\lambda y\, r]{\downarrow}$ (39.2), it follows by Rule =I (118.1) that $[\lambda y\, r] = [\lambda y\, r]$. Conjoining our assumption and this last fact, we obtain: $(r \equiv p) \,\&\, [\lambda y\, r] = [\lambda y\, r]$. By $\exists$I, it follows that $\exists q((q \equiv p) \,\&\, [\lambda y\, r] = [\lambda y\, q])$. So, by ($\vartheta$), it follows that $\circ p[\lambda y\, r]$. From this, it follows that $\circ p \Sigma r$, by (295). $\bowtie$

(299.4)$\star$ Since z is a variable substitutable for x in $\mathit{TruthValueOf}(x,p)$ and doesn't occur free in $\mathit{TruthValueOf}(x,p)$, we have the following instance of theorem (141)$\star$, where φ in that theorem is set to $\mathit{TruthValueOf}(x,p)$:

$$x = \imath x \mathit{TruthValueOf}(x,p) \equiv \forall z(\mathit{TruthValueOf}(z,p) \equiv z = x)$$

So by GEN:

$$(\vartheta)\quad \forall x(x = \imath x \mathit{TruthValueOf}(x,p) \equiv \forall z(\mathit{TruthValueOf}(z,p) \equiv z = x))$$

But by (293), $\imath x \mathit{TruthValueOf}(x,p){\downarrow}$. Hence by definition (294), we know $\circ p{\downarrow}$. From this last result and (ϑ), it follows by Rule $\forall$E (93.1) Variant that:

$$\circ p = \imath x \mathit{TruthValueOf}(x,p) \equiv \forall z(\mathit{TruthValueOf}(z,p) \equiv z = \circ p)$$

So, by definition (294) and biconditional syllogism, it follows that:

$$\forall z(\mathit{TruthValueOf}(z,p) \equiv z = \circ p)$$

By instantiating this last result to x, and we have:

$$\mathit{TruthValueOf}(x,p) \equiv x = \circ p \qquad \bowtie$$

(300)$\star$ Since $\circ p{\downarrow}$ and $\circ q{\downarrow}$, we can instantiate these into (288.3) to obtain:

$$(\mathit{TruthValueOf}(\circ p,p) \,\&\, \mathit{TruthValueOf}(\circ q,q)) \rightarrow (\circ p = \circ q \equiv (p \equiv q))$$

But by (299.1)$\star$, we know both $\mathit{TruthValueOf}(\circ p,p)$ and $\mathit{TruthValueOf}(\circ q,q)$. Hence $\circ p = \circ q \equiv (p \equiv q)$ $\bowtie$

(300)$\star$ [Alternative Proof] ($\rightarrow$) Assume $\circ p = \circ q$. Since we know that $\circ p \Sigma p$ by (296.2), it follows by Rule =E that $\circ q \Sigma p$. Hence by (299.3)$\star$, it follows that $p \equiv q$. ($\leftarrow$) Assume $p \equiv q$. We want to show that $\circ p = \circ q$. Clearly, $A!\circ p$ and $A!\circ q$. So by theorem (245.2), it remains only to show $\forall F(\circ pF \equiv \circ qF)$. Since F isn't free in our assumptions, it suffices by GEN to show $\circ pF \equiv \circ qF$. Without loss of generality, we prove only $\circ pF \rightarrow \circ qF$, since the reasoning for the converse is analogous. So assume $\circ pF$. From this and an alphabetic variant of (299.2)$\star$ it follows that:

$$\exists r((r \equiv p) \,\&\, F = [\lambda y\, r])$$

Let r_1 be an arbitrary such proposition, so that we know:

$$(\vartheta)\quad (r_1 \equiv p) \,\&\, F = [\lambda y\, r_1]$$

From the first conjunct of (ϑ) and our global assumption that $p \equiv q$, it follows that $r_1 \equiv q$. Conjoining this result with the second conjunct of (ϑ) and we have: $(r_1 \equiv q)\,\&\,F = [\lambda y\, r_1]$. So by ∃I, we have:

$$\exists r((r \equiv q)\,\&\,F = [\lambda y\, r])$$

From this it follows by (299.2)★ that $\circ q F$. ⋈

(301)★ By (299.1)★, we know $\mathit{TruthValueOf}(\circ q, q)$. Hence, $\exists p \mathit{TruthValueOf}(\circ q, p)$. So by definition (290), it follows that $\mathit{TruthValue}(\circ q)$. ⋈

(302.3) By (255) and the definitions of $\top$ (302.1) and $\bot$ (302.2), we know that both $\top$ and $\bot$ are abstract. So to show they are distinct, it suffices to show that there is a property one encodes that the other doesn't (245.3). Let p_0 be the necessary proposition $\forall x(E!x \rightarrow E!x)$ and consider the property $[\lambda z\, p_0]$, which exists by (39.2). Our proof strategy is to show that $\top[\lambda z\, p_0]$ and then show by reductio that $\neg\bot[\lambda z\, p_0]$. We first prove the following:

> *Lemma*: $\Box\exists p(p\,\&\,[\lambda z\, p_0] = [\lambda z\, p])$
>
> *Proof*: Clearly, p_0 is derivable from the tautology $E!x \rightarrow E!x$ and GEN, and $[\lambda z\, p_0] = [\lambda z\, p_0]$, by Rule =I (118.1) and the fact that $[\lambda z\, p_0]{\downarrow}$. So by conjoining the two and applying Rule ∃I, we have $\exists p(p\,\&\,[\lambda z\, p_0] = [\lambda z\, p])$. Since we established this by modally strict means from no assumptions, our lemma holds by RN.

Now if we let φ be $\exists p(p\,\&\,F = [\lambda z\, p])$, then our Lemma has the form $\Box\varphi_F^{[\lambda z\, p_0]}$. Note that by theorem (259.1), the definition of $\top$, and Rule =E, we know $\Box\varphi_F^G \rightarrow \top G$. Since this holds for any G, and we know the antecedent holds when G is $[\lambda z\, p_0]$, it follows that $\top[\lambda z\, p_0]$. Now assume $\bot[\lambda z\, p_0]$, for reductio. If we let ψ be $\exists p(\neg p\,\&\,F = [\lambda z\, p])$, then by theorem (258.2) and the definition of $\bot$, we know $\bot G \equiv \mathcal{A}\psi_F^G$. Since this holds for any G, it holds for $[\lambda z\, p_0]$. So our assumption implies $\mathcal{A}\psi_F^{[\lambda z\, p_0]}$, i.e., $\mathcal{A}\exists p(\neg p\,\&\,[\lambda z\, p_0] = [\lambda z\, p])$. Since the actuality operator commutes with the existential quantifier (139.10), it follows that $\exists p\mathcal{A}(\neg p\,\&\,[\lambda z\, p_0] = [\lambda z\, p])$. Let p_1 be such a proposition, so that we know $\mathcal{A}(\neg p_1\,\&\,[\lambda z\, p_0] = [\lambda z\, p_1])$. Since actuality distributes over a conjunction (139.2), we therefore know both:

(ϑ) $\mathcal{A}\neg p_1$

(ξ) $\mathcal{A}[\lambda y\, p_0] = [\lambda y\, p_1]$

It follows from (ξ) by (175.1) that $[\lambda y\, p_0] = [\lambda y\, p_1]$. Since p_0 and p_0 exist, it follows by the definition of identity for propositions (23.4) that $p_0 = p_1$, which by symmetry (117.2) yields $p_1 = p_0$. From this and (ϑ), we know $\mathcal{A}\neg p_0$, which by axiom (44.1) implies $\neg\mathcal{A}p_0$. But since p_0 is $\forall x(E!x \to E!x)$, $\Box p_0$ is a theorem, which implies $\mathcal{A}p_0$, by (132). Contradiction. ⋈

(303.1)★ Since the definiens of $\top$ (302.1) is canonical and thereby significant (252), it follows by the Rule of Identity by Definition (120) that:

$$(\vartheta)\quad \top = \imath x(A!x \,\&\, \forall F(xF \equiv \exists p(p \,\&\, F = [\lambda y\, p])))$$

So $\top\!\downarrow$, by (107.1). By applying GEN to an appropriate instance of (145.2)★ and instantiating $\top$, it follows that:

$$A!\top \,\&\, \forall F(\top F \equiv \exists p(p \,\&\, F = [\lambda y\, p]))$$

Hence by lemma (291.1), it follows that $TruthValue(\top)$. ⋈

(303.2)★ (Exercise)

(304.1)★ Assume $TruthValueOf(x,p)$. By definition (286), it follows that:

$$(\vartheta)\quad A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q]))$$

Independently, by the reasoning in the proof of (303.1)★, we know:

(a) $A!\top$

(b) $\forall F(\top F \equiv \exists q(q \,\&\, F = [\lambda y\, q]))$

Now we want to show $p \equiv x = \top$. ($\to$) Assume p. Since we know both x and $\top$ are abstract, it suffices by (245.2) to show that $\forall F(xF \equiv \top F)$. Note that since p is true, we know by theorem (289.1) that:

$$\forall F[\exists q(q \,\&\, F = [\lambda y\, q]) \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])]$$

By (99.10), it follows from (b) and this last result that:

$$\forall F(\top F \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q]))$$

By (99.11), we can commute the conditions of a quantified biconditional, to obtain:

(c) $\forall F(\exists q((q \equiv p) \,\&\, F = [\lambda y\, q]) \equiv \top F)$

Again by (99.10), the right conjunct of (ϑ) and (c) jointly imply $\forall F(xF \equiv \top F)$.

($\leftarrow$) Assume $x = \top$. Then by Rule =E, it follows from the right conjunct of (ϑ) that:

(d) $\forall F(\top F \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q]))$

By by commuting the biconditional in (b) (99.11) and applying (99.10) to the result and (d), it follows that:

$$\forall F[\exists q((q \equiv p) \,\&\, F = [\lambda y\, q]) \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])]$$

Hence by (289.1), it follows that p. ⋈

(304.2)★ (Exercise)

(305.1)★ By (299.1)★, we know $TruthValueOf(\circ p, p)$. So by the previous theorem (304.1)★, it follows that $p \equiv \circ p = \top$. ⋈

(305.2)★ (Exercise)

(305.3)★ (→) Assume p. Then by (305.1)★, $\circ p = \top$. But we know independently by (296.2), that $\circ p \Sigma p$. Hence by Rule =E, $\top \Sigma p$. (←) Assume $\top \Sigma p$. Then, by definition (295), $\top[\lambda y\, p]$. By the reasoning in (303.1)★, we already know that:

$$\forall F(\top F \equiv \exists q(q \,\&\, F = [\lambda y\, q])$$

Hence, $\exists q(q \,\&\, [\lambda y\, p] = [\lambda y\, q])$. Suppose q_1 is an arbitrary such proposition, so that we know both that q_1 is true and that $[\lambda y\, p] = [\lambda y\, q_1]$. Then by definition of proposition identity, $p = q_1$. Hence p is true. ⋈

(305.4)★ (Exercise)

(305.5)★ (Exercise)

(305.6)★ (Exercise)

(307) (→) Assume $ExtensionOf(x, p)$. So by definition (306):

$$A!x \,\&\, \forall F(xF \to Propositional(F)) \,\&\, \forall q((x\Sigma q) \equiv (q \equiv p))$$

And by definition (275) and the Rule of Substitution for Defined Formulas (160.3), it follows that:

$$(\vartheta) \quad A!x \,\&\, \forall F(xF \to \exists q(F = [\lambda y\, q])) \,\&\, \forall q((x\Sigma q) \equiv (q \equiv p))$$

Since the first conjunct is $A!x$, it remains by definition (286) to show:

$$\forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q]))$$

By GEN, it suffices to show:

$$xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])$$

We prove both directions:

(→) Assume xF. Then by the second conjunct of (ϑ), it follows that $\exists q(F = [\lambda y\, q])$. Let q_1 be an arbitrary such proposition, so that we know $F = [\lambda y\, q_1]$. Hence, $x[\lambda y\, q_1]$, and so it follows by definition (295) that $x\Sigma q_1$. But then from the third conjunct of (ϑ), it follows that $q_1 \equiv p$. So we've established $(q_1 \equiv p) \,\&\, F = [\lambda y\, q_1]$. Hence, $\exists q((q \equiv p) \,\&\, F = [\lambda y\, q])$.

($\leftarrow$) Assume $\exists q((q \equiv p) \,\&\, F = [\lambda y\, q])$. Let q_2 be an arbitrary such proposition, so that we know:

(ξ) $(q_2 \equiv p) \,\&\, F = [\lambda y\, q_2]$

Now by the third conjunct of (ϑ), it follows that $(x\Sigma q_2) \equiv (q_2 \equiv p)$. From this and the first conjunct of (ξ), it follows that $x\Sigma q_2$. By definition (295), it follows that $x[\lambda y\, q_2]$. But then by the second conjunct of (ξ), it follows that xF.

($\leftarrow$) Assume $\mathit{TruthValueOf}(x,p)$. So by definition (286), we know:

(ζ) $A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q]))$

Since $A!x$, it remains by definition (306) to show:

(a) $\forall F(xF \rightarrow \mathit{Propositional}(F))$

(b) $\forall q((x\Sigma q) \equiv (q \equiv p))$

(a) is easy, since it follows *a fortiori* from the second conjunct of (ζ) that $\forall F(xF \rightarrow \exists q(F = [\lambda y\, q]))$. This implies (a), by definition (275) and the Rule of Substitution for Defined Formulas (160.3). For (b), it suffices to show $(x\Sigma q_3) \equiv (q_3 \equiv p)$, for an arbitrary proposition q_3:

($\rightarrow$) Assume $x\Sigma q_3$. Then by definition (295), $x[\lambda y q_3]$. So by the second conjuct of (ζ), it follows that:

$$\exists q((q \equiv p) \,\&\, [\lambda y\, q_3] = [\lambda y\, q])$$

Suppose q_4 is an arbitrary such proposition, so that we know: $(q_4 \equiv p) \,\&\, [\lambda y\, q_3] = [\lambda y\, q_4]$. But the second conjunct of this last result and a well-known fact about the identity of propositions jointly imply $q_3 = q_4$. So from the first conjunct, it follows that $q_3 \equiv p$.

($\leftarrow$) Assume $q_3 \equiv p$. Then by Rule =I (118.2) and &I, $(q_3 \equiv p) \,\&\, [\lambda y q_3] = [\lambda y q_3]$. Hence, $\exists q((q \equiv p) \,\&\, [\lambda y\, q_3] = [\lambda y\, q])$. So by the second conjunct of (ζ), it follows that $x[\lambda y\, q_3]$. So by (295), $x\Sigma q_3$.

$\bowtie$

(308.1) By (288.2), $\exists! x\mathit{TruthValueOf}(x,p)$. Since (307) is modally strict, we may use it and the Rule of Substitution (160.2) to infer $\exists! x\mathit{ExtensionOf}(x,p)$. $\bowtie$

(308.2) (Exercise)

(308.3) As an instance of theorem (149.1), we know:

$$\mathcal{A}\forall x(\mathit{ExtensionOf}(x,p) \equiv \mathit{TruthValueOf}(x,p)) \rightarrow \\ \forall x(x = \iota x\mathit{ExtensionOf}(x,p) \equiv x = \iota x\mathit{TruthValueOf}(x,p))$$

But the antecedent is obtained by applying GEN and then the Rule of Actualization to theorem (307). Hence:

$$\forall x(x = \imath x TruthValueOf(x,p) \equiv x = \imath x ExtensionOf(x,p))$$

Instantiating to $\circ p$ (which we know exists), we obtain:

$$\circ p = \imath x TruthValueOf(x,p) \equiv \circ p = \imath x ExtensionOf(x,p)$$

But the left condition holds by definition (294). Hence $\circ p = \imath x ExtensionOf(x,p)$, which by the symmetry of identity gives us $\imath x ExtensionOf(x,p) = \circ p$. ⋈

(314.1) Assume $ExtensionOf(x,G)$ and $ExtensionOf(y,H)$. By definition (312.1), the fact that $G\downarrow$, and Rule $\equiv$S, these assumptions imply, respectively:

(a) $A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz))$

(b) $A!y \,\&\, \forall F(yF \equiv \forall z(Fz \equiv Hz))$

($\rightarrow$) Assume $x = y$. Then by Rule =E, it follows from (a) that:

(c) $A!y \,\&\, \forall F(yF \equiv \forall z(Fz \equiv Gz))$

Hence, by (99.11) and (99.10), the second conjuncts of (b) and (c) imply:

(d) $\forall F[\forall z(Fz \equiv Hz) \equiv \forall z(Fz \equiv Gz)]$

Now if we instantiate (d) to G, it follows that:

(e) $\forall z(Gz \equiv Hz) \equiv \forall z(Gz \equiv Gz)$

But the right condition of (e) is easily derivable. So $\forall z(Gz \equiv Hz)$, by biconditional syllogism.

($\leftarrow$) Assume:

(f) $\forall z(Gz \equiv Hz)$

Since we know $A!x$ and $A!y$ by the left conjuncts of (a) and (b), it suffices by theorem (245.2) to show $\forall F(xF \equiv yF)$, and by GEN, that $xF \equiv yF$:

> ($\rightarrow$) Assume xF. Then by the right conjunct of (a), it follows that $\forall z(Fz \equiv Gz)$. But from this and (f), it follows that $\forall z(Fz \equiv Hz)$. Hence, by the right conjunct of (b), it follows that yF.
>
> ($\leftarrow$) Assume yF. Then by the right conjunct of (b), it follows that $\forall z(Fz \equiv Hz)$. Now since our assumption (f), which asserts the material equivalence of G and H, is a symmetric condition on G and H (99.11), it follows that $\forall z(Hz \equiv Gz)$. But material equivalence is also a transitive condition on properties (99.10). Hence it follows from $\forall z(Fz \equiv Hz)$ and $\forall z(Hz \equiv Gz)$ that $\forall z(Fz \equiv Gz)$. So by the right conjunct of (a), xF. ⋈

(314.2) Assume $ExtensionOf(x,H))$, xF, and xG. From the first, it follows by definition (312.1) that:

(ϑ) $\forall K(xK \equiv \forall z(Kz \equiv Hz))$

From our assumptions xF, xG, and (ϑ), it follows respectively that $\forall z(Fz \equiv Hz)$ and $\forall z(Gz \equiv Hz)$. Hence $\forall z(Fz \equiv Gz)$. ⋈

(314.3) Assume xF, xG, and $\neg\forall z(Fz \equiv Gz)$. Assume, for reductio, that $Class(x)$. Then by definition, $\exists H(ExtensionOf(x,H))$. Suppose P is such a property, so that we know $ExtensionOf(x,P)$. Now, independently, by instantiating P into (314.2), we know that:

$$(ExtensionOf(x,P) \,\&\, xF \,\&\, xG) \to \forall z(Fz \equiv Gz)$$

Since we've established all three conjuncts of the antecedent, it follows that $\forall z(Fz \equiv Gz)$. Contradiction. ⋈

(314.3) [Alternative Proof] By classical propositional logic, (314.2) is equivalent to (exercise):

$$(xF \,\&\, xG \,\&\, \neg\forall z(Fz \equiv Gz)) \to \neg ExtensionOf(x,H)$$

By GEN, this holds universally for H, and since H doesn't occur free in the antecedent of the resulting universal claim, it follows by the right-to-left direction of (99.7) that:

$$(xF \,\&\, xG \,\&\, \neg\forall z(Fz \equiv Gz)) \to \forall H \neg ExtensionOf(x,H)$$

Hence, by the modally strict equivalence $\forall\alpha\neg\varphi \equiv \neg\exists\alpha\varphi$, it follows by a Rule of Substitution that:

$$(xF \,\&\, xG \,\&\, \neg\forall z(Fz \equiv Gz)) \to \neg\exists H(ExtensionOf(x,H))$$

Hence, by definition of $Class(x)$ (312.2) and a Rule of Substitution:

$$(xF \,\&\, xG \,\&\, \neg\forall z(Fz \equiv Gz)) \to \neg Class(x) \qquad ⋈$$

(315.1) As an instance of the Comprehension Principle for Abstract Objects (53), we know:

(ϑ) $\exists x(A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz)))$

But from definition (312.1) and the fact that $G{\downarrow}$, it becomes a modally strict theorem, by Rule $\equiv$S, that:

(ξ) $ExtensionOf(x,G) \equiv (A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz)))$

It then follows from (ϑ) and (ξ) that $\exists x ExtensionOf(x,G)$, by a Rule of Substitution (160.2). ⋈

(315.2) (Exercise)

(315.3) It follows from (315.2) that $\mathcal{A}\exists!x ExtensionOf(x,G)$, by the Rule of Actualization. Hence by (176.2), $\imath x ExtensionOf(x,G)\downarrow$. ⋈

(315.4) (Exercise)

(317.1) Assume $ExtensionOf(x,H)$. Then, by definition (312.1) it follows that:

$$\forall F(xF \equiv \forall z(Fz \equiv Hz))$$

We want to show $\forall y(y \in x \equiv Hy)$. Since y isn't free in our assumption, it suffices by GEN to show $y \in x \equiv Hy$.

($\rightarrow$) Assume $y \in x$. Then by definition of membership (316), it follows that $\exists G(ExtensionOf(x,G) \,\&\, Gy)$. Suppose P is an arbitrary such property, so that we know $ExtensionOf(x,P) \,\&\, Py$. Note independently that it follows from pre-Basic Law V (314.1) that:

$$(\vartheta) \quad (ExtensionOf(x,P) \,\&\, ExtensionOf(x,H)) \rightarrow (x=x \equiv \forall z(Pz \equiv Hz))$$

But we know both conjuncts of the antecedent. So we may infer from (ϑ) that $x = x \equiv \forall z(Pz \equiv Hz)$. By the reflexivity of identity, (117.1), it follows that $\forall z(Pz \equiv Hz)$. But since we also know Py, it follows that Hy.

($\leftarrow$) Assume Hy. But given our initial assumption, we have $ExtensionOf(x,H) \,\&$ Hy. So $\exists G(ExtensionOf(x,G) \,\&\, Gy)$. Hence, by definition of membership (316), $y \in x$. ⋈

(317.2) Assume, for reductio, $[\lambda x\, x \notin x]\downarrow$. Then by (315.1):

$$\exists y ExtensionOf(y,[\lambda x\, x \notin x])$$

Suppose a is such an object, so that we know $ExtensionOf(a,[\lambda x\, x \notin x])$. Then by (317.1), we know:

$$\forall y(y \in a \equiv [\lambda x\, x \notin x]y)$$

Instantiating to a, we therefore know:

$$a \in a \equiv [\lambda x\, x \notin x]a$$

However, since $[\lambda x\, x \notin x]$ exists, it follows from λ-Conversion that $[\lambda x\, x \notin x]a \equiv a \notin a$. So by biconditional syllogism, $a \in a \equiv a \notin a$. Contradiction. ⋈

(317.3) Assume, for reductio, $\exists x ExtensionOf(x,[\lambda x\, x \notin x])$ and that, say, a is such an object. Then it follows *a fortiori* from definition (312.1) that $[\lambda x x \notin x]\downarrow$, which contradicts (317.2). ⋈

(318) By (315.1), we know $\exists x ExtensionOf(x,F)$. Let a be an arbitrary such object, so that we know $ExtensionOf(a,F)$. Hence, $\exists G ExtensionOf(a,G)$, and so by definition (312.2), $Class(a)$. But also follows from $ExtensionOf(a,F)$, by (317.1), that $\forall y(y \in a \equiv Fy)$. Hence, we may conjoin our results to conclude $Class(a,F)\,\&\, \forall y(y \in a \equiv Fy)$. By $\exists$I, it follows that $\exists x(Class(x,F)\,\&\,\forall y(y \in x \equiv Fy))$. Since we derived this from no assumptions, it follows by GEN that $\forall F \exists x(Class(x,F)\,\&\, \forall y(y \in x \equiv Fy))$. ⋈

(319) By applying GEN to theorem (315.1) and instantiating to $A!$, we may infer $\exists x ClassOf(x,A!)$. Let b be such an object, so that we know $ClassOf(b,A!)$. It follows from this, *a fortiori* by definition (312), that $A!b$. Conjoining what we know, we have $ClassOf(b,A!)\,\&\,A!b$. Hence, $\exists G(ClassOf(b,G)\,\&\,Gb)$. So by definition (316), $b \in b$. But $ClassOf(b,A!)$ also implies $\exists F(ClassOf(x,F))$, and so by (315.2), $Class(b)$. Hence $Class(b)\,\&\,b \in b$. So $\exists x(Class(x)\,\&\,x \in x)$. ⋈

(320.1) By the usual principles of predicate logic, it suffices to show:

$$\exists x \exists G(ExtensionOf(x,G)\,\&\,\neg\Box ExtensionOf(x,G))$$

Before we look for witnesses to this claim, recall that by theorem (221.1), we know $\exists F \exists G(\forall z(Fz \equiv Gz)\,\&\,\Diamond\neg\forall z(Fz \equiv Gz))$. Let P and Q be such properties, so that we know:

$$(\vartheta)\quad \forall z(Pz \equiv Qz)\,\&\,\Diamond\neg\forall z(Pz \equiv Qz)$$

Now by (315.1), we know $\exists x ExtensionOf(x,Q)$. Let a be such an object, so that we know $ExtensionOf(a,Q)$. So it remains only to show $\neg\Box ExtensionOf(a,Q)$. Before we do, note that from $ExtensionOf(a,Q)$, we may infer $\forall F(aF \equiv \forall z(Fz \equiv Qz))$, by definition (312.1). This last fact and the first conjunct of (ϑ) yield aP.

Now suppose, for reductio, that $\Box ExtensionOf(a,Q)$. Then, by definition (312.1), and a Rule of Substitution, it follows that:

$$\Box(A!a\,\&\,Q{\downarrow}\,\&\,\forall F(aF \equiv \forall z(Fz \equiv Qz)))$$

By (158.3), all three conjuncts are necessary and so $\Box\forall F(aF \equiv \forall z(Fz \equiv Qz))$, in particular. Then by the Converse Barcan Formula (167.2):

$$\forall F \Box(aF \equiv \forall z(Fz \equiv Qz))$$

Instantiating to P and we obtain: $\Box(aP \equiv \forall z(Pz \equiv Qz))$. So by (158.6), $\Box aP \equiv \Box\forall z(Pz \equiv Qz)$. But we established above that aP, and so by axiom (51), $\Box aP$. Hence $\Box\forall z(Pz \equiv Qz)$, i.e., $\neg\Diamond\neg\forall z(Pz \equiv Qz)$, which contradicts the second conjunct of (ϑ). ⋈

(320.1) [Alternative Proof] By the usual principles of predicate logic, it suffices to show:

$$\exists x \exists G(ExtensionOf(x,G)\,\&\,\neg\Box ExtensionOf(x,G))$$

Before we look for a witness to this claim, recall that by theorem (221.1), we know $\exists F \exists G(\forall z(Fz \equiv Gz) \,\&\, \Diamond\neg\forall z(Fz \equiv Gz))$. Let P and Q be such properties, so that we know:

(ϑ) $\forall z(Pz \equiv Qz) \,\&\, \Diamond\neg\forall z(Pz \equiv Qz)$

Now by (315.1), we know $\exists x ExtensionOf(x, Q)$. Let a be such an object, so that we know $ExtensionOf(a, Q)$. So it remains only to show $\neg\Box ExtensionOf(a, Q)$. Note that from $ExtensionOf(a, Q)$, it follows *a fortiori*, from definition (312.1), that:

$$\forall F(aF \equiv \forall z(Fz \equiv Qz))$$

This last fact implies not only aQ, but together with the first conjunct of (ϑ), implies aP. So by axiom (51), $\Box aP$ and $\Box aQ$. Now since we want to show $\neg\Box ExtensionOf(a, Q)$, assume, for reductio, that $\Box ExtensionOf(a, Q)$. Then we know:

$$\Box ExtensionOf(a, Q) \,\&\, \Box aP \,\&\, \Box aQ$$

Hence, by the right-to-left direction of an extended version of (158.3):

(ζ) $\Box(ExtensionOf(a, Q) \,\&\, aP \,\&\, aQ)$

Now, independently, if we apply RN and then GEN to the modally strict theorem (314.2), we obtain:

$$\forall H \forall F \forall G \forall x \Box((ExtensionOf(x, H)) \,\&\, xF \,\&\, xG) \to \forall z(Fz \equiv Gz))$$

Instantiating to Q, P, Q, and a, respectively:

$$\Box((ExtensionOf(a, Q)) \,\&\, aP \,\&\, aQ) \to \forall z(Pz \equiv Qz))$$

So by the K axiom:

$$\Box(ExtensionOf(a, Q)) \,\&\, aP \,\&\, aQ) \to \Box\forall z(Pz \equiv Qz)$$

But the antecedent is just (ζ). Hence $\Box\forall z(Pz \equiv Qz)$, i.e., $\neg\Diamond\neg\forall z(Pz \equiv Qz)$, which contradicts the second conjunct of (ϑ). ⋈

(320.2) By the usual principles of predicate logic, it suffices to show:

$$\exists x(Class(x) \,\&\, \neg\Box Class(x))$$

Before we identify a witness, recall that $\exists F \exists G(\forall x(Fx \equiv Gx) \,\&\, \Diamond\neg\forall x(Fx \equiv Gx))$, by (221.1). Suppose P and Q are such properties, so that we know:

(A) $\forall z(Pz \equiv Qz)$

(B) $\Diamond\neg\forall z(Pz \equiv Qz)$

Now by (315.1), we know $\exists x ExtensionOf(x,Q)$. So let a be such an object, so that we know:

(C) $ExtensionOf(a,Q)$

Since (C) implies $\exists H(ExtensionOf(a,H))$, we may infer by definition (312.2) that $Class(a)$. It then remains to show $\neg\Box Class(a)$, and so it suffices to show $\Diamond\neg Class(a)$. Our strategy is to use (314.3), which by RN and GEN implies:

$$\forall x\forall F\forall G\Box((xF \,\&\, xG \,\&\, \neg\forall z(Fz \equiv Gz)) \to \neg Class(x)) \qquad (314.3)$$

But first note that (C) also implies, by definition (312.1), that:

(D) $\forall F(aF \equiv \forall z(Fz \equiv Qz))$

By now familiar reasoning, we may infer from (D) and (A) that aP and from (D) and $\forall z(Qz \equiv Qz)$ that aQ. So by axiom (51):

(E1) $\Box aP$

(E2) $\Box aQ$

Now if we instantiate (314.3) to a, P, and Q, it follows that:

$$\Box((aP \,\&\, aQ \,\&\, \neg\forall z(Pz \equiv Qz)) \to \neg Class(a))$$

From this it follows by the K$\Diamond$ principle (158.13) that:

$$\Diamond(aP \,\&\, aQ \,\&\, \neg\forall z(Pz \equiv Qz)) \to \Diamond\neg Class(a)$$

So if we can establish the antecedent, we're done. But note that the following is an easy theorem of K:

$$(\Box\varphi \,\&\, \Box\psi \,\&\, \Diamond\chi) \to \Diamond(\varphi \,\&\, \psi \,\&\, \chi)$$

This is a variant of (158.16) and we leave it as an exercise. As an instance, we know:

$$(\Box aP \,\&\, \Box aQ \,\&\, \Diamond\neg\forall z(Pz \equiv Qz)) \to \Diamond(aP \,\&\, aQ \,\&\, \neg\forall z(Pz \equiv Qz))$$

But by &I, we've established the antecedent, since $\Box aP$ is (E1), $\Box aQ$ is (E2), and $\Diamond\neg\forall z(Pz \equiv Qx)$ is (B). $\bowtie$

(320.3) To show our theorem, it suffices to show $\Diamond\neg\forall x(Class(x) \to \mathcal{A}Class(x))$, i.e., to show $\Diamond\exists x(Class(x) \,\&\, \neg\mathcal{A}Class(x))$. By CBF$\Diamond$ (167.4), it suffices to show:

(ϑ) $\exists x\Diamond(Class(x) \,\&\, \neg\mathcal{A}Class(x))$

So we have to find a witness to (ϑ). Before we do so, we first prove an instrumental, modally strict lemma:

(ξ) $\forall F \forall G \Box(\forall z(Fz \equiv Gz) \to \exists x(Class(x) \,\&\, xF \,\&\, xG))$

Proof. By GEN, we show $\Box(\forall z(Fz \equiv Gz) \to \exists x(Class(x) \,\&\, xF \,\&\, xG))$. By RN, it suffices to find a modally strict proof of $\forall z(Fz \equiv Gz) \to \exists x(Class(x) \,\&\, xF \,\&\, xG)$. So assume $\forall z(Fz \equiv Gz)$. Now we know independently by (315.1) $\exists x ExtensionOf(x, G)$. So let a be such an object, so that we know $ExtensionOf(a, G)$. It then remains only to show that a is a witness to $\exists x(Class(x) \,\&\, xF \,\&\, xG)$. Since $\exists H(ExtensionOf(a, H))$, it follows by (312.2), that $Class(a)$. It also follows from $ExtensionOf(a, G)$, by definition (312.1), that $\forall H(aH \equiv \forall z(Hz \equiv Gz))$. But, our assumption is $\forall z(Fz \equiv Gz)$, and so it also follows that aF. And, clearly, aG.

Now to prove our theorem, note that by (221.3), we know $\exists F \exists G(\mathcal{A}\neg\forall z(Fz \equiv Gz) \,\&\, \Diamond\forall z(Fz \equiv Gz))$. So let P and Q be such properties, so that we know:

(A) $\mathcal{A}\neg\forall z(Pz \equiv Qz)$

(B) $\Diamond\forall z(Pz \equiv Qz)$

Moreover, instantiating P and Q into our lemma (ξ), we know that the following is a theorem:

$$\Box(\forall z(Pz \equiv Qz) \to \exists x(Class(x) \,\&\, xP \,\&\, xQ))$$

Hence, by the K$\Diamond$ schema, $\Diamond\forall z(Pz \equiv Qz) \to \Diamond\exists x(Class(x) \,\&\, xP \,\&\, xQ)$. From this last result and (B), it follows that $\Diamond\exists x(Class(x) \,\&\, xP \,\&\, xQ)$. So by CBF (167.3), $\exists x\Diamond(Class(x) \,\&\, xP \,\&\, xQ)$. Now suppose b is such an object, so that we know:

$$\Diamond(Class(b) \,\&\, bP \,\&\, bQ)$$

This implies, by (162.3):

(ζ) $\Diamond Class(b) \,\&\, \Diamond bP \,\&\, \Diamond bQ$

Now if we can show that b is the desired witness to (ϑ), we're done. So we have to show $\Diamond(Class(b) \,\&\, \neg\mathcal{A}Class(b))$. But this follows from (158.16) if we can show: $\Diamond Class(b) \,\&\, \Box\neg\mathcal{A}Class(b)$. So, by &I, it suffices to show both:

(C) $\Diamond Class(b)$

(D) $\Box\neg\mathcal{A}Class(b)$

But (C) is the first conjunct of (ζ). So it remains to show (D). Now if we apply the Rule of Actualization (RA) and then GEN thrice to theorem (314.3), then we know:

$$\forall F \forall G \forall x \mathcal{A}((xF \,\&\, xG \,\&\, \neg\forall z(Fz \equiv Gz)) \to \neg Class(x))$$

Instantiating to P, Q, and b:

$$\mathcal{A}((bP \,\&\, bQ \,\&\, \neg\forall z(Pz \equiv Qz)) \to \neg Class(b))$$

By applications of axiom (44.2) and theorem (139.2), it follows that:

$$(\mathcal{A}bP \,\&\, \mathcal{A}bQ \,\&\, \mathcal{A}\neg\forall z(Pz \equiv Qz)) \to \mathcal{A}\neg Class(b)$$

The antecedent of this last result holds, since:

- the first conjunct follows from the 2nd conjunct of (ζ), by (179.3) and (179.10),
- the second conjunct follows from the 3rd conjunct of (ζ), by (179.3) and (179.10), and
- the third conjunct is (A).

Thus $\mathcal{A}\neg Class(b)$. So by axiom (46.1), $\Box\mathcal{A}\neg Class(b)$, which is all that remained for us to show. ⋈

(320.4) Assume $ExtensionOf(x, H)$. Then it follows from definition (312.1) that:

(ϑ) $\forall F(xF \equiv \forall z(Fz \equiv Hz))$

Hence, xH, and by axiom (51), $\Box xH$. Now by (221.4), it follows that for some property G, H is equivalent to G but might not have been, i.e., that $\exists G(\forall z(Hz \equiv Gz) \,\&\, \Diamond\neg\forall z(Hz \equiv Gz))$. Suppose Q is such a property, so that we know both $\forall z(Hz \equiv Qz)$ and $\Diamond\neg\forall z(Hz \equiv Qz)$. It follows from the first and (ϑ) that xQ, and so $\Box xQ$. Now as an instance of (314.2), substituting H for H, H for F, and Q for G, we know:

$$(ExtensionOf(x, H) \,\&\, xH \,\&\, xQ) \to \forall z(Hz \equiv Qz)$$

By classical propositional logic (exercise), it follows that:

$$(xH \,\&\, xQ \,\&\, \neg\forall z(Hz \equiv Qz)) \to \neg ExtensionOf(x, H)$$

Since this is a modally strict theorem, it follows by RM$\Diamond$ that:

(ξ) $\Diamond(xH \,\&\, xQ \,\&\, \neg\forall z(Hz \equiv Qz)) \to \Diamond\neg ExtensionOf(x, H)$

But we already know all of $\Box xH$, $\Box xQ$ and $\Diamond\neg\forall z(Hz \equiv Qz)$. But we saw in the proof of (320.2) that by reasoning from (158.16), these three facts imply $\Diamond(xH \,\&\, xQ \,\&\, \neg\forall z(Hz \equiv Qz))$. So by ($\xi$), $\Diamond\neg ExtensionOf(x, H)$. ⋈

(320.5) Assume $Class(x)$. Then by definition (312.2), $\exists F(ClassOf(x, F))$. Let P be such a property, so that we know $ClassOf(x, P)$. By definition (312.1), it follows that:

(ϑ) $\forall F(xF \equiv \forall z(Fz \equiv Pz))$

Hence, xP, and by axiom (51), $\Box xP$. Independently, if we instantiate P into theorem (221.4), then we know there is a property G such that P is equivalent to G but possibly not, i.e., that $\exists G(\forall z(Pz \equiv Gz) \,\&\, \Diamond\neg\forall z(Pz \equiv Gz))$. Let Q be such a property, so that we know:

(ξ) $\forall z(Pz \equiv Qz) \,\&\, \Diamond\neg\forall z(Pz \equiv Qz)$

From (ϑ) and the first conjunct of (ξ), we may infer xQ, and so $\Box xQ$, again by (51). But from the second conjunct of (ξ) and the facts that $\Box xP$ and $\Box xQ$, it follows by now familiar modal reasoning based on (158.16) that (exercise):

$$\Diamond(xP \,\&\, xQ \,\&\, \neg\forall z(Pz \equiv Qz))$$

But if we apply RM$\Diamond$ to theorem (314.3) and instantiate to P and Q, we also know:

$$\Diamond(xP \,\&\, xQ \,\&\, \neg\forall z(Pz \equiv Qz)) \rightarrow \Diamond\neg Class(x)$$

Hence $\Diamond\neg Class(x)$. ⋈

(321.1) By standard principles of predicate logic, it suffices to establish:

$$\exists x \exists y(y \in x \,\&\, \neg\Box y \in x)$$

We start with the fact that $\exists p ContingentlyTrue(p)$ (217.1). Suppose p_1 is such a proposition, so that we know $p_1 \,\&\, \Diamond\neg p_1$. Then consider $[\lambda x\, p_1]$, which clearly exists. Call this property Q. By (315.1), we know $\exists x ExtensionOf(x, Q)$. Let a be such an object, so that we know $ExtensionOf(a, Q)$. By definition (312), it follows that:

(A) $\forall F(aF \equiv \forall z(Fz \equiv Qz))$

And by β-Conversion, we know, where y is an arbitrary but fixed object:

(B) $Qy \equiv p_1$

Now consider the property we've previously designated as L, i.e., $[\lambda x\, E!x \rightarrow E!x]$. Then by the reasoning in the proof of (221.1), we know that L and Q are materially equivalent but possibly not materially equivalent:

(C) $\forall z(Lz \equiv Qz) \,\&\, \Diamond\neg\forall z(Lz \equiv Qz)$

From (A), (B), and (C) it is straightforward to establish aQ, aL, and Qy. From the first two, it follows by axiom (51) that $\Box aQ$ and $\Box aL$. From the third, we now know $ExtensionOf(a, Q) \,\&\, Qy$, and so it follows by $\exists$I and definition (316) that $y \in a$. So it remains only to show $\neg\Box y \in a$.

First note that by the principles of propositional logic, theorem (314.2) is equivalent to (exercise):

$$(xF \,\&\, xG \,\&\, \neg\forall z(Fz \equiv Gz)) \to \neg \mathit{ExtensionOf}(x,H)$$

If we apply GEN to universally generalize on H, then by the right to left direction of (39.7) it follows that:

$$(xF \,\&\, xG \,\&\, \neg\forall z(Fz \equiv Gz)) \to \forall H \neg \mathit{ExtensionOf}(x,H)$$

Hence, by the modally strict equivalence $\forall\alpha\neg\varphi \equiv \neg\exists\alpha\varphi$, it follows by a Rule of Substitution that:

$$(xF \,\&\, xG \,\&\, \neg\forall z(Fz \equiv Gz)) \to \neg\exists H(\mathit{ExtensionOf}(x,H))$$

Since this is a modally strict theorem, it follows by RM$\Diamond$ that:

$$\Diamond(xF \,\&\, xG \,\&\, \neg\forall z(Fz \equiv Gz)) \to \Diamond\neg\exists H(\mathit{ExtensionOf}(x,H))$$

By GEN, this holds for any x, F, and G, and so it holds for a, L, and Q:

(D) $\Diamond(aL \,\&\, aQ \,\&\, \neg\forall z(Lz \equiv Qz)) \to \Diamond\neg\exists H(\mathit{ExtensionOf}(x,H))$

But by now familiar reasoning from (158.16), the antecedent of (D) follows from $\Box aL$, $\Box aQ$, and $\Diamond\neg\forall z(Lz \equiv Qz)$, all of which we've established. Hence $\Diamond\neg\exists H(\mathit{ExtensionOf}(a,H))$. But $\Diamond\neg\exists\alpha\varphi \to \Diamond\neg\exists\alpha(\varphi \,\&\, \psi)$ is also a modal principle – this is easily established as a modal consequence of the contrapositive of (103.5). So it follows that $\Diamond\neg\exists H(\mathit{ExtensionOf}(a,H) \,\&\, Hy)$. Hence, by definition (316) and a Rule of Substitution (160.3), it follows that $\Diamond\neg y \in a$, i.e., $\neg\Box y \in a$. ⋈

(321.2) Assume $y \in x$. Then $\exists F(\mathit{ExtensionOf}(x,F) \,\&\, Fy)$. Suppose P is such a property, so that we know both $\mathit{ExtensionOf}(x,P)$ and Py. From the first it follows *a fortiori* from definition (312.1) that:

(ϑ) $\forall F(xF \equiv \forall z(Fz \equiv Pz))$

Hence, xP, and so $\Box xP$. Now by (221.4), we independently know that, for some property G, P is equivalent to G but might not have been. So suppose Q is such a property, so that we know both $\forall x(Px \equiv Qx)$ and $\Diamond\neg\forall x(Px \equiv Qx)$. The first and ($\vartheta$) imply xQ, and so $\Box xQ$. But by now familiar reasoning from (158.16), $\Box xP$, $\Box xQ$ and $\Diamond\neg\forall x(Px \equiv Qx)$ imply:

$$\Diamond(xP \,\&\, xQ \,\&\, \neg\forall x(Px \equiv Qx)$$

But by applying RM$\Diamond$ to the relevant instance of (314.3), we know:

$$\Diamond(xP \,\&\, xQ \,\&\, \neg\forall x(Px \equiv Qx) \to \Diamond\neg \mathit{Class}(x)$$

Hence, $\Diamond\neg\mathit{Class}(x)$, and so by definition (312.2) and a Rule of Substitution, $\Diamond\neg\exists F(\mathit{ExtensionOf}(x,F))$. It follows *a fortiori*, that $\Diamond\neg\exists F(\mathit{ExtensionOf}(x,F) \& Fy)$, by reasoning from a modal consequence of the contrapositive of (103.5), which we also used at the end of (321.1). Hence $\Diamond\neg y \in x$, i.e., $\neg\Box y \in x$. ⋈

(323.1) Let Π be any unary relation term in which x doesn't occur free. Then, to prove our theorem, we first establish the following Lemma:

Lemma: $\neg\Pi\downarrow \vdash \neg\exists!x ExtensionOf(x,\Pi)$

Proof. Assume $\neg\Pi\downarrow$. Then given that $\neg\varphi \rightarrow \neg(\psi \,\&\, \varphi \,\&\, \chi)$, we know:

$$\neg(A!x \,\&\, \Pi\downarrow \,\&\, \forall F(xF \equiv \forall z(Fz \equiv \Pi z)))$$

Hence, by definition of $ExtensionOf(x,G)$ (312.1), $\neg ExtensionOf(x,\Pi)$. By hypothesis, x isn't free in Π, and so it isn't free in our assumption. By GEN it follows that $\forall x \neg ExtensionOf(x,\Pi)$, i.e., $\neg\exists x ExtensionOf(x,\Pi)$. So by the definition of the uniqueness quantifier, $\neg\exists!x ExtensionOf(x,\Pi)$. Since we've established $\neg\Pi\downarrow \rightarrow \neg\exists!x ExtensionOf(x,\Pi)$ by conditional proof, our Lemma follows by (63.10).

We then prove our theorem as follows: Assume $\neg\Pi\downarrow$. Then by (106.2), it follows that $\Box\neg\Pi\downarrow$. So by (132), $\mathcal{A}\neg\Pi\downarrow$. But it follows from our Lemma by Rule RA that: $\mathcal{A}\neg\Pi\downarrow \vdash \mathcal{A}\neg\exists!x ExtensionOf(x,\Pi)$. Hence $\mathcal{A}\neg\exists!x ExtensionOf(x,\Pi)$. So by axiom (44.1), $\neg\mathcal{A}\exists!x ExtensionOf(x,\Pi)$. Hence, by (176.1):

$$\neg\exists!x \mathcal{A} ExtensionOf(x,\Pi)$$

So by (176.2), $\neg\imath x ExtensionOf(x,\Pi)\downarrow$. ⋈

(323.2) Assume $\neg\Pi\downarrow$. Without loss of generality, pick some variable, say x, that doesn't occur free in Π. Then by (323.1), we have $\neg\imath x ExtensionOf(x,\Pi)\downarrow$. Now the definition of ϵG (322) and the Rule of Definition by Identity (73), we know that:

$$(\imath x ExtensionOf(x,\Pi)\downarrow \rightarrow \epsilon\Pi = \imath x ExtensionOf(x,\Pi)) \,\&\, \\ (\neg\imath x ExtensionOf(x,\Pi)\downarrow \rightarrow \neg\epsilon\Pi\downarrow)$$

It therefore follows from $\neg\imath x ExtensionOf(x,\Pi)\downarrow$ and the second conjunct that $\neg\epsilon\Pi\downarrow$. ⋈

(324.1) By (315.3), we know:

(A) $\imath x ExtensionOf(x,G)\downarrow$

So, by definition (322) and the Rule of Identity by Definition (120.1) :

(B) $\epsilon G = \imath x ExtensionOf(x,G)$

Now, independently, definition (312.1) and the $\vdash_\Box$ version of the Rule of Equivalence by Definition (90.1) jointly imply that the following is a modally strict theorem:

$$ExtensionOf(x,G) \equiv A!x \,\&\, G\downarrow \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz))$$

By Rule $\equiv$S of Biconditional Simplification, the following is therefore a modally strict theorem:

$$ExtensionOf(x,G) \equiv A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz))$$

So by GEN and RN:

(C) $\Box\forall x(ExtensionOf(x,G) \equiv A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz)))$

It follows from (A) and (C) by theorem (149.3) that:

(D) $\imath xExtensionOf(x,G) = \imath x(A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz)))$

By the 2nd Exercise in (117), we know that the transitivity of identity can be applied to (B) and (D) to conclude:

$$\epsilon G = \imath x(A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz))) \qquad \bowtie$$

(324.2) Since the tautology $Gz \equiv Gz$ and GEN yield the easy theorem that $\forall z(Gz \equiv Gz)$, it follows by the Rule of Actualization that $\mathcal{A}\forall z(Gz \equiv Gz)$. If we let φ be $\forall z(Fz \equiv Gz)$, then we have established $\mathcal{A}\varphi^G_F$. But by (258.2), we know:

$$\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G \equiv \mathcal{A}\varphi^G_F$$

Hence $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G$. But since φ is $\forall z(Fz \equiv Gz)$, it follows by (324.1) and Rule =E that ϵGG. $\bowtie$

(325) Theorem (221.1) is $\exists F\exists G(\forall z(Fz \equiv Gz) \,\&\, \Diamond\neg\forall z(Fz \equiv Gz))$. Let P and Q be such properties, so that we know:

$$\forall z(Pz \equiv Qz) \,\&\, \Diamond\neg\forall z(Pz \equiv Qz)$$

Hence, by the T$\Diamond$ schema (163.1):

$$\Diamond(\forall z(Pz \equiv Qz) \,\&\, \Diamond\neg\forall z(Pz \equiv Qz))$$

So by two applications of $\exists$I:

$$\exists G\exists F\Diamond(\forall z(Fz \equiv Gz) \,\&\, \Diamond\neg\forall z(Fz \equiv Gz)) \qquad \bowtie$$

(327.1)★ By definition (322), we know $\epsilon G = \imath xExtensionOf(x,G)$. This in turn implies $\epsilon G\downarrow$. So by (145.2)★, $ExtensionOf(\epsilon G,G)$. $\bowtie$

(327.2)★ By (327.1)★, $ExtensionOf(\epsilon G,G)$. When we expand this by the definition of *ExtensionOf* (312.1), our theorem, $\forall F(\epsilon GF \equiv \forall z(Fz \equiv Gz))$, is the third conjunct. $\bowtie$

(327.3)★ By theorem (141)★ and GEN, we know:

(ϑ) $\forall x(x = \imath xExtensionOf(x,G) \equiv \forall z(ExtensionOf(z,G) \equiv z = x))$

But by (315.3), $\imath x ExtensionOf(x,G)\downarrow$. So by definition (322), $\epsilon G\downarrow$. If we instantiate (ϑ) to ϵG, we obtain:

$$\epsilon G = \imath x ExtensionOf(x,G) \equiv \forall z(ExtensionOf(z,G) \equiv z = \epsilon G)$$

So, by definition (322) and biconditional syllogism, it follows that:

$$\forall z(ExtensionOf(z,G) \equiv z = \epsilon G)$$

By instantiating this last result to x, and we have:

$$ExtensionOf(x,G) \equiv x = \epsilon G \qquad \bowtie$$

(328)★ Since $\epsilon F\downarrow$ and $\epsilon G\downarrow$ exist, we can instantiate ϵF and ϵG into pre-Basic Law V (314.1). Substituting ϵF for x, F for G, ϵG for y, and G for H, we obtain:

$$(ExtensionOf(\epsilon F,F) \,\&\, ExtensionOf(\epsilon G,G)) \to (\epsilon F = \epsilon G \equiv \forall z(Fz \equiv Gz))$$

But we also know by (327.1)★ that $ExtensionOf(\epsilon F,F)$ and $ExtensionOf(\epsilon G,G)$. Hence, $\epsilon F = \epsilon G \equiv \forall z(Fz \equiv Gz)$. $\bowtie$[460]

(331.1)[461] Suppose, for reductio, that $[\lambda x\, \exists G(x = \epsilon G \,\&\, \neg Gx)]\downarrow$. Let's *abbreviate* the λ-expression in this claim as K, so that our reductio assumption becomes $K\downarrow$. Then by (315.3), $\imath z ExtensionOf(z,K)\downarrow$, and by definition (322), $\epsilon K\downarrow$. We can derive a contradiction with the help of the following lemma, namely, that for any property F, ϵK encodes F if and only if it is actually the case that F is materially equivalent to K:

Lemma: $\forall F(\epsilon KF \equiv \mathcal{A}\forall z(Fz \equiv Kz))$

Proof: By GEN, we prove $\epsilon KF \equiv \mathcal{A}\forall z(Fz \equiv Kz)$. As an instance of (258.1), we know: $\imath x(A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Kz)))F \equiv \mathcal{A}\forall z(Fz \equiv Az)$. But, by (324.1), we know $\epsilon K = \imath x(A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Kz)))$. So by Rule =E, $\epsilon KF \equiv \mathcal{A}\forall z(Fz \equiv Kz))$.

---

[460] Alternatively: ($\to$) Assume $\epsilon F = \epsilon G$. By (327.2)★, we know that $\forall H(\epsilon FH \equiv \forall x(Hx \equiv Fx))$. Since $\epsilon F = \epsilon G$, it follows by Rule =E that $\forall H(\epsilon GH \equiv \forall x(Hx \equiv Fx))$. Instantiating the universal claim to G, it follows that $\epsilon GG \equiv \forall x(Gx \equiv Fx)$. Since ϵGG by (324.2), it follows that $\forall x(Gx \equiv Fx)$, which in turn implies $\forall x(Fx \equiv Gx)$. ($\leftarrow$) Take $\forall x(Fx \equiv Gx)$ as a global assumption. By (324.1), ϵF and ϵG are canonical individuals, and so by (255), they are both abstract. So to show they are identical it suffices, by (245.2), to show they encode the same properties. By GEN, it suffices to show $\epsilon FH \equiv \epsilon GH$:

($\to$) Assume ϵFH. Then by the left-to-right direction of (327.2)★, $\forall x(Hx \equiv Fx)$. From this and our global assumption it follows that $\forall x(Hx \equiv Gx)$. But by the right-to-left direction of (327.2)★, it follows that ϵGH.

($\leftarrow$) By analogous reasoning. $\bowtie$

[461] I'm indebted to Daniel West, who pointed out and proved that this theorem and the next are modally strict. This improved an earlier version of this monograph, in which I had given only non-modally strict proofs. The following argument follows his proof in most of the details.

With this Lemma, we can reason to the contradiction $\mathcal{A}K\epsilon K \equiv \neg\mathcal{A}K\epsilon K$, i.e., it is actually the case that ϵK exemplifies K if and only if it is not actually the case that ϵK exemplifies K.

($\rightarrow$) Assume $\mathcal{A}K\epsilon K$. As an instance of Rule $\overrightarrow{\beta}$C (184.1.a), we know the following, since the formula to the left of $\vdash$ abbreviates $[\lambda x\, \exists G(x = \epsilon G \,\&\, \neg Gx)]\epsilon K$:

$$K\epsilon K \vdash \exists G(\epsilon K = \epsilon G \,\&\, \neg G\epsilon K)$$

So by the Rule of Actualization (135):

$$\mathcal{A}K\epsilon K \vdash \mathcal{A}\exists G(\epsilon K = \epsilon G \,\&\, \neg G\epsilon K)$$

Given our assumption, it follows that $\mathcal{A}\exists G(\epsilon K = \epsilon G \,\&\, \neg G\epsilon K)$. So by theorem (139.10), $\exists G\mathcal{A}(\epsilon K = \epsilon G \,\&\, \neg G\epsilon K)$. Suppose P is such a property, so that we know $\mathcal{A}(\epsilon K = \epsilon P \,\&\, \neg P\epsilon K)$. By (139.2) and &E, we know both:

(ϑ) $\mathcal{A}(\epsilon K = \epsilon P)$

(ξ) $\mathcal{A}\neg P\epsilon K$

But (ϑ) implies $\epsilon K = \epsilon P$, by (175.1). Since we know independently by (324.2) that ϵPP, it follows that ϵKP. So by the above Lemma, it follows that $\mathcal{A}\forall z(Pz \equiv Kz)$. Hence, by axiom (44.3), $\forall z\mathcal{A}(Pz \equiv Kz)$. Instantiating this latter to ϵK, it follows that $\mathcal{A}(P\epsilon K \equiv K\epsilon K)$. So by theorem (139.5), $\mathcal{A}P\epsilon K \equiv \mathcal{A}K\epsilon K$. But it follows from ($\xi$) that $\neg\mathcal{A}P\epsilon K$, by axiom (44.1). Hence $\neg\mathcal{A}K\epsilon K$, by biconditional syllogism.

($\leftarrow$) Assume $\neg\mathcal{A}K\epsilon K$. As an instance of Rule $\overleftarrow{\beta}$C (184.2.b), we know the following:

$$K\downarrow,\ \epsilon K\downarrow,\ \neg K\epsilon K \vdash \neg\exists G(\epsilon K = \epsilon G \,\&\, \neg G\epsilon K)$$

So by the Rule of Actualization:

(ζ) $\mathcal{A}K\downarrow,\ \mathcal{A}\epsilon K\downarrow,\ \mathcal{A}\neg K\epsilon K \vdash \mathcal{A}\neg\exists G(\epsilon K = \epsilon G \,\&\, \neg G\epsilon K)$

But $\mathcal{A}K\downarrow$ follows from: our assumption $K\downarrow$, the instance $K\downarrow \rightarrow \Box K\downarrow$ of theorem (106), and the instance $\Box K\downarrow \rightarrow \mathcal{A}K\downarrow$ of theorem (132). By analogous reasoning, $\mathcal{A}\epsilon K\downarrow$. And our assumption implies $\mathcal{A}\neg K\epsilon K$, by axiom (44.1). From these facts and (ζ), we may conclude $\mathcal{A}\neg\exists G(\epsilon K = \epsilon G \,\&\, \neg G\epsilon K)$. Now it is a modally strict theorem that $\neg\exists\alpha(\varphi \,\&\, \neg\psi) \equiv \forall\alpha(\varphi \rightarrow \psi)$ (exercise). So it follows from our last result by an appropriate instance of this theorem and a Rule of Substitution that $\mathcal{A}\forall G(\epsilon K = \epsilon G \rightarrow G\epsilon K)$. Hence $\forall G\mathcal{A}(\epsilon K = \epsilon G \rightarrow G\epsilon K)$. If we instantiate this to K, it follows that $\mathcal{A}(\epsilon K = \epsilon K \rightarrow K\epsilon K)$. So $\mathcal{A}\epsilon K = \epsilon K \rightarrow \mathcal{A}K\epsilon K$, by distributing the $\mathcal{A}$ operator. Since we already $\epsilon K\downarrow$, we can invoke Rule =I (118.1) to establish $\epsilon K = \epsilon K$. So by (175.1), $\mathcal{A}\epsilon K = \epsilon K$. Hence $\mathcal{A}K\epsilon K$. Contradiction. ⋈

(331.2) By (331.1), $\neg[\lambda x\, \exists G(x = \epsilon G \,\&\, \neg Gx)]\downarrow$. So $\neg\epsilon[\lambda x\, \exists G(x = \epsilon G \,\&\, \neg Gx)]\downarrow$, by (323.2). ⋈

(333)★ (Exercise)

(334)★ It follows from (327.1)★ that $\exists G(\mathit{ExtensionOf}(\epsilon G, G))$ and so by theorem (312.2), $\mathit{Class}(\epsilon G)$. ⋈

(341.3.a) We want to show that the following metarule is justified:

If $\Gamma(\gamma_1,\ldots,\gamma_n) \vdash_\Box \varphi(\gamma_1,\ldots,\gamma_n)$, then $\Box\Gamma(\gamma_1,\ldots,\gamma_n) \vdash \Box\varphi(\gamma_1,\ldots,\gamma_n)$

Our strategy is to show that when we apply our conventions for restricted variables, this abbreviates a metarule justified by RN. Now to see what our metatheorem abbreviates, give the conjunctive interpretation to each formula in $\Gamma(\gamma_1,\ldots,\gamma_n)$ and $\Box\Gamma(\gamma_1,\ldots,\gamma_n)$, and the give the conditional interpretation to $\varphi(\gamma_1,\ldots,\gamma_n)$ and $\Box\varphi(\gamma_1,\ldots,\gamma_n)$. As noted in the discussion in the text, the latter is easy:

$\varphi(\gamma_1,\ldots,\gamma_n)$ abbreviates $(\psi_1 \,\&\, \ldots \,\&\, \psi_n) \to \varphi$
$\Box\varphi(\gamma_1,\ldots,\gamma_n)$ abbreviates $(\psi_1 \,\&\, \ldots \,\&\, \psi_n) \to \Box\varphi$

However, to give the conjunctive interpretation to each formula in $\Gamma(\gamma_i,\ldots,\gamma_j)$, we follow the indication in footnote 234:

interpret each $\chi(\gamma_1,\ldots,\gamma_n)$ in $\Gamma(\gamma_1,\ldots,\gamma_n)$ as $\psi_1 \,\&\, \ldots \,\&\, \psi_n \,\&\, \chi$
interpret each $\Box\chi(\gamma_1,\ldots,\gamma_n)$ in $\Box\Gamma(\gamma_1,\ldots,\gamma_n)$ as $\psi_1 \,\&\, \ldots \,\&\, \psi_n \,\&\, \Box\chi$

Now we leave it as an exercise to show that the following equivalences govern our strict derivability relation:

$\psi_1 \,\&\, \ldots \,\&\, \psi_n \,\&\, \chi \vdash_\Box \varphi$ if and only if $\psi_1,\ldots,\psi_n,\chi \vdash_\Box \varphi$

$\Gamma \vdash_\Box (\psi_1 \,\&\, \ldots \,\&\, \psi_n) \to \varphi$ if and only if $\psi_1,\ldots,\psi_n,\Gamma \vdash_\Box \varphi$

Hence we may validly *factor out* all of the ψ_is (a) from all the premises in Γ and $\Box\Gamma$ and (b) from the conclusions $(\psi_1 \,\&\, \ldots \,\&\, \psi_n) \to \varphi$ and $(\psi_1 \,\&\, \ldots \,\&\, \psi_n) \to \Box\varphi$. It therefore suffices to show that the following metarule is justified:

If $\psi_1,\ldots,\psi_n,\Gamma \vdash_\Box \varphi$, then $\psi_1,\ldots,\psi_n,\Box\Gamma \vdash_\Box \Box\varphi$, where $\psi_1,\ldots,\psi_n$ are rigid restriction conditions.

So assume:

$\psi_1,\ldots,\psi_n,\Gamma \vdash_\Box \varphi$

We want to show:

$\psi_1,\ldots,\psi_n,\Box\Gamma \vdash_\Box \Box\varphi$

Note first that it follows from our assumption by RN that:

$\Box\psi_1, \ldots, \Box\psi_n, \Box\Gamma \vdash_{\Box} \Box\varphi$

Now we leave it as an exercise to show that $\chi_1, \ldots, \chi_n, \Gamma \vdash_{\Box} \varphi$ if and only if $\chi_1 \mathbin{\&} \ldots \mathbin{\&} \chi_n, \Gamma \vdash_{\Box} \varphi$. Given this principle, it follows from our last displayed result that:

(ϑ) $\Box\psi_1 \mathbin{\&} \ldots \mathbin{\&} \Box\psi_n, \Box\Gamma \vdash_{\Box} \Box\varphi$

Independently, since $\psi_1, \ldots, \psi_n$ are, by hypothesis, rigid restriction conditions, we know all of the following:

$\vdash_{\Box} \forall\alpha(\psi_1 \to \Box\psi_1)$
$\vdots$
$\vdash_{\Box} \forall\alpha(\psi_n \to \Box\psi_n)$

Hence, by $\forall$E:

$\vdash_{\Box} \psi_1 \to \Box\psi_1$
$\vdots$
$\vdash_{\Box} \psi_n \to \Box\psi_n$

So by (63.10), we know:

$\psi_1 \vdash_{\Box} \Box\psi_1$
$\vdots$
$\psi_n \vdash_{\Box} \Box\psi_n$

By enough applications of (86.1), it follows from this last sequence of results that:

(ξ) $\psi_1, \ldots, \psi_n \vdash_{\Box} \Box\psi_1 \mathbin{\&} \ldots \mathbin{\&} \Box\psi_n$

Thus, from (ξ) and (ϑ) it follows by (63.8) that:

$\psi_1, \ldots, \psi_n, \Box\Gamma \vdash_{\Box} \Box\varphi$ ⋈

(341.3.b) By the reasoning at the outset of (341.3.a), it suffices to show that the following metarule is justified:

If $\psi_1, \ldots, \psi_n, \Gamma \vdash \varphi$, then $\psi_1, \ldots, \psi_n, \mathcal{A}\Gamma \vdash \mathcal{A}\varphi$, where $\psi_1, \ldots, \psi_n$ are rigid restriction conditions.

So assume:

$\psi_1, \ldots, \psi_n, \Gamma \vdash \varphi$

We want to show:

$$\psi_1, \ldots, \psi_n, \mathcal{A}\Gamma \vdash \mathcal{A}\varphi$$

Note first that it follows from our assumption by RA that:

$$\mathcal{A}\psi_1, \ldots, \mathcal{A}\psi_n, \mathcal{A}\Gamma \vdash \mathcal{A}\varphi$$

Now we leave it as an exercise to show that $\chi_1, \ldots, \chi_n, \Gamma \vdash \varphi$ if and only if $\chi_1 \mathbin{\&} \ldots \mathbin{\&} \chi_n, \Gamma \vdash \varphi$. Given this principle, it follows from our last displayed result that:

$$(\vartheta) \quad \mathcal{A}\psi_1 \mathbin{\&} \ldots \mathbin{\&} \mathcal{A}\psi_n, \mathcal{A}\Gamma \vdash \mathcal{A}\varphi$$

Independently, since (a) $\psi_1, \ldots, \psi_n$ are, by hypothesis, rigid restriction conditions, and (b) modally strict theorems are theorems (62.2), we know all of the following:

$$\begin{array}{l} \vdash \forall\alpha(\psi_1 \to \Box\psi_1) \\ \qquad \vdots \\ \vdash \forall\alpha(\psi_n \to \Box\psi_n) \end{array}$$

Hence, by $\forall$E:

$$\begin{array}{l} \vdash \psi_1 \to \Box\psi_1 \\ \qquad \vdots \\ \vdash \psi_n \to \Box\psi_n \end{array}$$

So by (63.10), we know:

$$\begin{array}{l} \psi_1 \vdash \Box\psi_1 \\ \qquad \vdots \\ \psi_n \vdash \Box\psi_n \end{array}$$

By enough applications of (86.1), it follows from this last sequence of results that:

$$(\xi) \quad \psi_1, \ldots, \psi_n \vdash \Box\psi_1 \mathbin{\&} \ldots \mathbin{\&} \Box\psi_n$$

It is also easy to establish, by theorem (132) and the rules for &, that (exercise):

$$(\zeta) \quad \Box\psi_1 \mathbin{\&} \ldots \mathbin{\&} \Box\psi_n \vdash \mathcal{A}\psi_1 \mathbin{\&} \ldots \mathbin{\&} \mathcal{A}\psi_n$$

From (ξ) and (ζ), it follows that (exercise):

$$\psi_1, \ldots, \psi_n \vdash \mathcal{A}\psi_1 \mathbin{\&} \ldots \mathbin{\&} \mathcal{A}\psi_n$$

Thus, from this last result and (ϑ) it follows by (63.8) that:

$$\psi_1, \ldots, \psi_n, \mathcal{A}\Gamma \vdash \mathcal{A}\varphi \qquad \bowtie$$

(343) If we eliminate all the restricted variables, then the theorem to be proved is:

$$\forall x \forall y((Class(x) \,\&\, Class(y)) \rightarrow (x = y \equiv \forall z(z \in x \equiv z \in y)))$$

So by GEN, it suffices to show:

$$(Class(x) \,\&\, Class(y)) \rightarrow (x = y \equiv \forall z(z \in x \equiv z \in y))$$

Assume $Class(x) \,\&\, Class(y)$. ($\rightarrow$) Exercise. ($\leftarrow$) Assume:

(a) $\forall z(z \in x \equiv z \in y)$

By definition (312.2), $Class(x)$ and $Class(y)$ imply $\exists F(ExtensionOf(x, F))$ and $\exists G(ExtensionOf(y, G))$. Let P and Q be such properties, so that $ExtensionOf(x, P)$ and $ExtensionOf(y, Q)$, respectively. But by theorem (317.1), it follows from the first that $\forall z(z \in x \equiv Pz)$, which is equivalent to:

(b) $\forall z(Pz \equiv z \in x)$

And it follows from the second that:

(c) $\forall z(z \in y \equiv Qy)$

From (b) and (a), it follows by (99.10) that $\forall z(Pz \equiv z \in y)$, and from this and (c) it follows, also by (99.10), that $\forall z(Pz \equiv Qz)$. Now, independently, the following is an instance of pre-Basic Law V (314):

$$(ExtensionOf(x, P) \,\&\, ExtensionOf(y, Q)) \rightarrow (x = y \equiv \forall z(Pz \equiv Qz))$$

Since both conjuncts of the antecedent are known, it follows that:

$$x = y \equiv \forall z(Pz \equiv Qz)$$

Since we've established the right side of the above biconditional, it follows that $x = y$. $\bowtie$

(345.1) By our convention (337.2) for bound restricted variables, we have to show:

(ϑ) $\exists x(Class(x) \,\&\, Empty(x))$

Moreover, by our convention (337.1) for definitions-by-equivalence with free restricted variables, (344) is an abbreviation of:

$$Empty(x) \equiv_{df} Class(x) \,\&\, \neg\exists y(y \in x)$$

By a Rule of Substitution (160.3), we can exchange the definiens and definiendum when they occur as subformulas, so to show (ϑ) it suffices to show:

$$\exists x(Class(x) \,\&\, Class(x) \,\&\, \neg\exists y(y \in x))$$

But given the otiose conjunct, it suffices by the idempotence of & (85.6) to show:

(ξ) $\exists x(Class(x) \,\&\, \neg\exists y(y \in x))$

Consider the impossible property $\overline{L}$, where L was defined as $[\lambda x\, E!x \rightarrow E!x]$ (203) and $\overline{F}$ was defined as $[\lambda y\, \neg Fy]$ (196). Then by the Fundamental Theorem for Natural Classes and Logical Sets (318), we know:

$$\exists x(Class(x) \,\&\, \forall y(y \in x \equiv \overline{L}y))$$

Suppose a is an arbitrary such object, so that we know:

(ζ) $Class(a) \,\&\, \forall y(y \in a \equiv \overline{L}y)$

It now suffices to show that a is a witness to (ξ). And given the first conjunct of (ζ), it remains to show $\neg\exists y(y \in a)$. Suppose, for reductio, that $\exists y(y \in a)$. Let b be an arbitrary such object, so that we know $b \in a$. Then by the second conjunct of (ζ), it follows that $\overline{L}b$. But recall that in the proof of (203.2), we developed a (modally-strict) proof of $\forall x \neg\overline{L}x$. So $\neg\overline{L}b$. Contradiction. ⋈

(345.2) Eliminating the restricted variables, we have to show:

$$\exists x(Class(x) \,\&\, Empty(x) \,\&\, \forall z((Class(z) \,\&\, Empty(z)) \rightarrow z = x))$$

By (345.1), we already know $\exists x(Class(x) \,\&\, Empty(x))$. So suppose a is an arbitrary such object, so that $Class(a) \& Empty(a)$. So to conclude that a is the needed witness, it remains to show $\forall z((Class(z) \,\&\, Empty(z)) \rightarrow z = a)$. Since z isn't free in our assumption, it suffices by GEN to show $(Class(z) \,\&\, Empty(z)) \rightarrow z = a$. So assume $Class(z) \,\&\, Empty(z)$. By our conventions for free restricted variables in definition-by-$\equiv$ (338.1), definition (344) implies:

$$Empty(a) \equiv Class(a) \,\&\, \neg\exists y(y \in a)$$

$$Empty(z) \equiv Class(z) \,\&\, \neg\exists y(y \in z)$$

Given what we've established, it follows that $\neg\exists y(y \in a)$ and $\neg\exists y(y \in z)$. So by (103.9), $\forall x(x \in z \equiv x \in a)$, i.e., that z and a have the same members. Hence, by the principle of extensionality (343), it follows that $z = a$. ⋈

(345.3) By (176.2), we know:

(ϑ) $\imath x(Class(x) \,\&\, Empty(x))\downarrow \equiv \mathcal{A}\exists! x(Class(x) \,\&\, Empty(x))$

But theorem (345.2) is $\exists!cEmpty(c)$, i.e., $\exists!x(Class(x) \& Empty(x))$. So by the Rule of Actualization, $\mathcal{A}\exists!x(Class(x) \& Empty(x))$. From this and ($\vartheta$), it follows that $\imath x(Class(x) \& Empty(x))\downarrow$. So by our conventions for restricted variables bound by term-forming operators (337.3), $\imath cEmpty(c)\downarrow$. ⋈

(347.1) (Exercise)

(347.2) By eliminating the restricted variable in definition (346), we know:

$$\varnothing = \imath x(Class(x) \& Empty(x))$$

Since all the terms involved are significant, it suffices, by the transitivity of identity, to show:

$$\imath x(Class(x) \& Empty(x)) = \imath x(A!x \& \forall F(xF \equiv \neg\exists zFz))$$

We show this by way of theorem (149.3). Given theorem (345.3), it remains only to establish:

$$(\vartheta)\quad \Box\forall x((Class(x) \& Empty(x)) \equiv (A!x \& \forall F(xF \equiv \neg\exists zFz)))$$

By GEN and RN, it suffices to show:

$$(Class(x) \& Empty(x)) \equiv (A!x \& \forall F(xF \equiv \neg\exists zFz))$$

($\rightarrow$) Assume both $Class(x)$ and $Empty(x)$. The latter implies by definition (344) that $\neg\exists y(y \in x)$. The former implies $\exists G(ExtensionOf(x, G))$, by (312.2). Suppose P is such a property, so that we know $ExtensionOf(x, P)$. Hence, by definition (312.1), we know both $A!x$ and:

$$(\zeta)\quad \forall F(xF \equiv \forall z(Fz \equiv Pz))$$

And $ExtensionOf(x, P)$ also implies, by (317.1):

$$(\xi)\quad \forall y(y \in x \equiv Py)$$

Now since we've established $A!x$, it remains to show $\forall F(xF \equiv \neg\exists zFz)$. By GEN, we show $xF \equiv \neg\exists zFz$. ($\rightarrow$) Assume xF. Then by (ζ), $\forall z(Fz \equiv Pz)$. Independently, from our hypothesis that $\neg\exists y(y \in x)$ and (ξ) it follows that $\neg\exists yPy$. But we've established that F is materially equivalent to P. So $\neg\exists zFz$. ($\leftarrow$) Assume $\neg\exists zFz$. But we know, $\neg\exists y(y \in x)$. Then by (ξ), $\neg\exists yPy$. Since F and P are both unexemplified, they are materially equivalent (103.9), i.e.,$\forall z(Fz \equiv Pz)$. Hence xF, by (ζ).

($\leftarrow$) Assume $A!x \& \forall F(xF \equiv \neg\exists zFz)$. We have to show:

$$\text{(A)}\quad Class(x)$$

and, by definition (344), show:

(B) $\neg\exists y(y \in x)$

(A) By (312.2), we want to show that $\exists G(\mathit{ExtensionOf}(x,G))$. To find a witness, consider the necessarily unexemplified property $\overline{L}$, which we defined previously as the negation (196) of L, which in turn was defined as $[\lambda x\, E!x \to E!x]$ (203). By definition (312.1) and $\exists$I, we need only show:

(C) $A!x \,\&\, \overline{L}{\downarrow} \,\&\, \forall F(xF \equiv \forall z(Fz \equiv \overline{L}z))$

But $A!x$ is already known. $\overline{L}{\downarrow}$ is also known, given the definition of $\overline{L}$. Now given our assumption $\forall F(xF \equiv \neg\exists zFz)$, to show the third conjunct of (C) we need only show $\forall F(\neg\exists zFz \equiv \forall z(Fz \equiv \overline{L}z))$, by properties of quantified biconditionals (99.10). So, by GEN, we show $\neg\exists zFz \equiv \forall z(Fz \equiv \overline{L}z)$. This is easy:

($\to$) Assume $\neg\exists zFz$. From this and the fact (exercise) that $\neg\exists z\overline{L}z$, it follows by (103.9) that $\forall z(Fz \equiv \overline{L}z)$.

($\leftarrow$) Assume $\forall z(Fz \equiv \overline{L}z)$. For reductio, assume $\exists zFz$ and suppose a is such an individual, so that we know Fa. Then since our local assumption implies $Fa \equiv \overline{L}a$, it follows that $\overline{L}a$, which contradicts the fact that $\neg\exists z\overline{L}z$. ⋈

(B) In the proof of (A) we established that $\mathit{ExtensionOf}(x,\overline{L})$. This implies, by (317.1), that:

(ξ) $\forall y(y \in x \equiv \overline{L}y)$

But $\overline{L}$ is an impossible property, and so we know $\neg\exists y\overline{L}y$. Hence, $\neg\exists y(y \in x)$. ⋈

(349.1) Consider the necessary property L, which we defined previously as $[\lambda x\, E!x \to E!x]$ (203). Then by the Fundamental Theorem for Natural Classes and Logical Sets (318), we know:

$$\exists x(\mathit{Class}(x) \,\&\, \forall y(y \in x \equiv Ly)$$

Suppose a is an arbitrary such class, so that we know:

(ϑ) $\forall y(y \in a \equiv Ly)$

So if a is to be our witness, it remains, by the definition of a universal* class (348), only to show $\forall y(y \in a)$. By GEN, we show $y \in a$. Since (ϑ) implies $y \in a \equiv Ly$, it suffices to show Ly. But by the definition of L, β-Conversion, and the fact that $L{\downarrow}$, we know $Ly \equiv (E!y \to E!y)$. Since the right side of this biconditional is a tautology, Ly. ⋈

(349.2) (Exercise)

(351.1) By our conventions for restricted variables, we have to show:

$$\forall x \forall y((\mathit{Class}(x) \,\&\, \mathit{Class}(y)) \to \exists z(\mathit{Class}(z) \,\&\, \mathit{UnionOf}(z,x,y)))$$

So by GEN, we have to show:

$$(Class(x) \,\&\, Class(y)) \rightarrow \exists z(Class(z) \,\&\, UnionOf(z,x,y))$$

So assume $Class(x) \,\&\, Class(y)$. By definition of *UnionOf* (350) and eliminating the otiose conjuncts from the definiens, we have to show:

$$\exists z(Class(z) \,\&\, \forall w(w \in z \equiv (w \in x \,\vee\, w \in y))$$

To find our witness, note that since x and y are classes, it follows by (312.2) that there are properties P and Q such that $ExtensionOf(x,P)$ and $ExtensionOf(y,Q)$. So by (317.1), we know both:

(ϑ) $\forall w(w \in x \equiv Pw)$

(ξ) $\forall w(w \in y \equiv Qw)$

Now consider the property $[\lambda z\, Pz \vee Qz]$. By (39.2), $[\lambda z\, Pz \vee Qz]\!\downarrow$. Then by the Fundamental Theorem for Natural Classes and Logical Sets (318), we know:

$$\exists x(Class(x) \,\&\, \forall y(y \in x \equiv [\lambda z\, Pz \vee Qz]y))$$

Let b be an arbitrary such object, so that we know $Class(b)$ and:

(ζ) $\forall y(y \in b \equiv [\lambda z\, Pz \vee Qz]y)$

Now if we can show $\forall w(w \in b \equiv (w \in x \vee w \in y))$, then b is the desired witness and we're done. By GEN, it suffices to show $w \in b \equiv (w \in x \vee w \in y)$. This is established by the following chain of biconditionals, all of which are consequences of what we have established so far:

$$\begin{array}{rcll} w \in b & \equiv & [\lambda z\, Pz \vee Qz]w & \text{by } (\zeta) \\ & \equiv & Pw \vee Qw & \text{by } [\lambda z\, Pz \vee Qz]\!\downarrow \text{ and Rule } \overrightarrow{\beta}\text{C } (184.1.\text{a}) \\ & \equiv & w \in x \vee Qw & \text{by } (\vartheta) \text{ and } (88.8.\text{h}) \\ & \equiv & w \in x \vee w \in y & \text{by } (\xi), \text{ and } (88.8.\text{g}) \quad \bowtie \end{array}$$

(351.2) By our conventions for restricted variables and GEN, we have to show:

$$(Class(x) \,\&\, Class(y)) \rightarrow \\ \exists z(Class(z) \& UnionOf(z,x,y) \& \forall w(Class(w) \& UnionOf(w,x,y) \rightarrow w = z))$$

So assume $Class(x) \,\&\, Class(y)$. We may therefore instantiate x and y into by (351.1), and conclude:

$$\exists z(Class(z) \,\&\, UnionOf(z,x,y))$$

Assume a is an arbitrary such class, so that $Class(a)$ and $UnionOf(a,x,y)$. By GEN, &I, and $\exists$I, it remains to show $Class(w) \,\&\, UnionOf(w,x,y) \rightarrow w = a$. So assume both $Class(w)$ and $UnionOf(w,x,y)$. Then, by definition of *UnionOf* (350), it follows that:

(ϑ) $\forall z(z \in w \equiv (z \in x \vee z \in y))$

Now since $UnionOf(a,x,y)$, we know by definition (350) that:

(ξ) $\forall z(z \in a \equiv (z \in x \vee z \in y))$

So by the logic of quantified biconditionals, (ϑ) and (ξ) jointly imply $\forall z(z \in w \equiv z \in a)$. It follows by the principle of extensionality (343) that $w = a$. ⋈

(352.1)$\star$ By eliminating the restricted variables and applying the Rule of Actualization to theorem (351.2), it follows that:

$$\mathcal{A}\forall x\forall y((Class(x) \,\&\, Class(y)) \to \exists! z(Class(z) \,\&\, UnionOf(z,x,y)))$$

Note that this is equivalent to:

(ϑ) $\mathcal{A}\forall x(Class(x) \to \forall y(Class(y) \to \exists! z(Class(z) \,\&\, UnionOf(z,x,y))))$

Now in footnote 231, in the discussion in Remark (340), we saw that it is a $\star$-theorem that $\mathcal{A}\forall\alpha(\psi \to \varphi) \to \forall\alpha(\psi \to \mathcal{A}\varphi)$. Hence ($\vartheta$) implies:

$$\forall x(Class(x) \to \mathcal{A}\forall y(Class(y) \to \exists! z(Class(z) \,\&\, UnionOf(z,x,y))))$$

So by $\forall$E:

$$Class(x) \to \mathcal{A}\forall y(Class(y) \to \exists! z(Class(z) \,\&\, UnionOf(z,x,y)))$$

Now, for conditional proof, assume $Class(x)$. Hence:

$$\mathcal{A}\forall y(Class(y) \to \exists! z(Class(z) \,\&\, UnionOf(z,x,y)))$$

We now repeat this sequence of reasoning. By the $\star$-theorem in footnote 231:

$$\forall y(Class(y) \to \mathcal{A}\exists! z(Class(z) \,\&\, UnionOf(z,x,y)))$$

And, again by $\forall$E:

$$Class(y) \to \mathcal{A}\exists! z(Class(z) \,\&\, UnionOf(z,x,y))$$

And, again, for conditional proof, assume $Class(y)$. Hence:

$$\mathcal{A}\exists! z(Class(z) \,\&\, UnionOf(z,x,y)))$$

So, by theorem (176.2), it follows that $\imath z(Class(z) \,\&\, UnionOf(z,x,y))\downarrow$. Hence, by conditional proof, we've established:

$$(Class(x) \,\&\, Class(y)) \to \imath z(Class(z) \,\&\, UnionOf(z,x,y))\downarrow$$

Since x and y aren't free in any undischarged assumption, it follows by GEN that:

$$\forall x\forall y((Class(x) \,\&\, Class(y)) \to \imath z(Class(z) \,\&\, UnionOf(z,x,y))\downarrow)$$

So by our conventions for restricted variables, $\forall c'\forall c''(\imath c UnionOf(c,c',c'')\downarrow)$. ⋈

(352.2)★ If we eliminate the restricted variables, we have to show:

$$\neg\Box\forall x\forall y((Class(x) \,\&\, Class(y)) \rightarrow \imath z(Class(z) \,\&\, UnionOf(z,x,y))\downarrow)$$

Suppose this were necessary, for reductio. Then by CBF (167.2):

$$(\vartheta)\ \ \forall x\forall y\Box((Class(x) \,\&\, Class(y)) \rightarrow \imath z(Class(z) \,\&\, UnionOf(z,x,y))\downarrow)$$

Now independently, note that by standard modal reasoning, theorem (320.3) is equivalent to:

$$\Diamond\exists x(Class(x) \,\&\, \neg\mathcal{A}Class(x))$$

So by BF$\Diamond$ (167.3):

$$\exists x\Diamond(Class(x) \,\&\, \neg\mathcal{A}Class(x))$$

Now suppose a is such an object, so that we know:

$$\Diamond(Class(a) \,\&\, \neg\mathcal{A}Class(a))$$

Hence, by (162.3):

$$(\xi)\ \ \Diamond Class(a) \,\&\, \Diamond\neg\mathcal{A}Class(a)$$

But since $\neg\mathcal{A}\varphi \equiv \mathcal{A}\neg\varphi$ is a necessary axiom (44.1) (and so a modally strict theorem), we can apply a Rule of Substitution to the second conjunct of (ξ) to infer $\Diamond\mathcal{A}\neg Class(a)$. Hence by (164.4), $\mathcal{A}\neg Class(a)$, and so by (43)★, $\neg Class(a)$. But this last result implies (exercise):

$$\neg\exists z(Class(z) \,\&\, Class(a) \,\&\, Class(a) \,\&\, \forall y(y \in z \equiv (y \in a \vee y \in a)))$$

So by adding an otiose conjunct, it follows, by definition (350) and a Rule of Substitution:

$$\neg\exists z(Class(z) \,\&\, UnionOf(z,a,a))$$

Hence, $\neg\exists!z(Class(z) \,\&\, UnionOf(z,a,a))$ and so by (144.1)★:

$$\neg\imath z(Class(z) \,\&\, UnionOf(z,a,a))\downarrow$$

But if we instantiate a twice into (ϑ), we obtain:

$$\Box((Class(a) \,\&\, Class(a)) \rightarrow \imath z(Class(z) \,\&\, UnionOf(z,a,a))\downarrow)$$

By the K$\Diamond$ principle, this implies:

$$(\zeta)\ \ \Diamond(Class(a) \,\&\, Class(a)) \rightarrow \Diamond\imath z(Class(z) \,\&\, UnionOf(z,a,a))\downarrow$$

But $\Diamond\varphi \to \Diamond(\varphi \,\&\, \varphi)$ (exercise), and so the first conjunct of (ζ) follows from the first conjunct of (ξ). Hence, $\Diamond\imath z(Class(z) \,\&\, UnionOf(z,a,a))\!\downarrow$. But by theorem (169.1), $\Diamond\tau\!\downarrow \to \tau\!\downarrow$. Hence $\imath z(Class(z) \,\&\, UnionOf(z,a,a))\!\downarrow$. Contradiction. ⋈

(353) To prove our theorem, we begin as we did in the previous proof, eliminating the restricted variables. Thus, we have to show:

$$\neg\Box\forall x\forall y((Class(x) \,\&\, Class(y)) \to \imath z(Class(z) \,\&\, UnionOf(z,x,y))\!\downarrow)$$

Suppose this were necessary, for reductio. Then by CBF (167.2):

$$(\vartheta)\quad \forall x\forall y\Box((Class(x) \,\&\, Class(y)) \to \imath z(Class(z) \,\&\, UnionOf(z,x,y))\!\downarrow)$$

Now independently, note that by standard modal reasoning, theorem (320.3) is equivalent to:

$$\Diamond\exists x(Class(x) \,\&\, \neg\mathcal{A}Class(x))$$

So by BF$\Diamond$ (167.3):

$$\exists x\Diamond(Class(x) \,\&\, \neg\mathcal{A}Class(x))$$

Now suppose a is such an object, so that we know:

$$\Diamond(Class(a) \,\&\, \neg\mathcal{A}Class(a))$$

Hence, by (162.3):

$$(\xi)\quad \Diamond Class(a) \,\&\, \Diamond\neg\mathcal{A}Class(a)$$

But the second conjunct of (ξ) is equivalent to $\Diamond\mathcal{A}\neg Class(a)$, by the necessary axiom (44.1) and a Rule of Substitution. Hence by (164.4), $\mathcal{A}\neg Class(a)$, and so by axiom (44.1), $\neg\mathcal{A}Class(a)$. Hence (exercise):

$$\neg\exists z\mathcal{A}(Class(z) \,\&\, Class(a) \,\&\, Class(a) \,\&\, \forall y(y \in z \equiv (y \in a \vee y \in a)))$$

So by adding an otiose conjunct and applying definition (350) and a Rule of Substitution:

$$\neg\exists z\mathcal{A}(Class(z) \,\&\, UnionOf(z,a,a))$$

A fortiori,

$$\neg\exists! z\mathcal{A}(Class(z) \,\&\, UnionOf(z,a,a))$$

So by theorem (152.1):

$$\neg\imath z(Class(z) \,\&\, UnionOf(z,a,a))\!\downarrow$$

But if we instantiate a twice into (ϑ), we obtain:

$$\Box((Class(a) \,\&\, Class(a)) \to \imath z(Class(z) \,\&\, UnionOf(z,a,a))\downarrow)$$

By the K$\Diamond$ principle, this implies:

$$(\zeta)\ \ \Diamond(Class(a) \,\&\, Class(a)) \to \Diamond\imath z(Class(z) \,\&\, UnionOf(z,a,a))\downarrow$$

But $\Diamond\varphi \to \Diamond(\varphi \,\&\, \varphi)$ (exercise), and so the first conjunct of (ζ) follows from the first conjunct of (ξ). Hence, $\Diamond\imath z(Class(z) \,\&\, UnionOf(z,a,a))\downarrow$. But by theorem (169.1), $\Diamond\tau\downarrow \to \tau\downarrow$. Hence $\imath z(Class(z) \,\&\, UnionOf(z,a,a))\downarrow$. Contradiction. ⋈

(355)★ We want to show:

$$\forall x\forall y((Class(x) \,\&\, Class(y)) \to \forall z(z \in x \cup y \equiv (z \in x \vee z \in y)))$$

So by GEN, it suffices to show:

$$(Class(x) \,\&\, Class(y)) \to \forall z(z \in x \cup y \equiv (z \in x \vee z \in y))$$

Assume $Class(x) \,\&\, Class(y)$. By our conventions for free restricted variables in definitions-by-= (339.1), the definition of $\cup$ (354) abbreviates:

$$(\vartheta)\ \ x \cup y =_{df} \imath z(Class(z) \,\&\, Class(x) \,\&\, Class(y) \,\&\, UnionOf(z,x,y))$$

Independently, since $Class(x) \,\&\, Class(y)$, it follows from (352.1)★ that:

$$(\xi)\ \ \imath z(Class(z) \,\&\, UnionOf(z,x,y))\downarrow$$

Similarly, by our conventions for free restricted variables in definitions-by-$\equiv$ (338.2), the definition of *UnionOf* (350) abbreviates:

$$UnionOf(z,x,y) \equiv_{df} Class(z) \& Class(x) \& Class(y) \& \forall w(w \in z \equiv w \in x \vee w \in y)$$

Since the definiens of $UnionOf(z,x,y)$ includes the conjuncts $Class(x)$ and $Class(y)$, it follows that (exercise):

$$\Box\forall z((Class(z) \& UnionOf(z,x,y)) \equiv (Class(z) \& Class(x) \& Class(y) \& UnionOf(z,x,y)))$$

From this last result and (ξ), it follows by (149.3) that:

$$\imath z(Class(z) \,\&\, Class(x) \,\&\, Class(y) \,\&\, UnionOf(z,x,y))\downarrow$$

So by definition (ϑ) and the theories of definition and identity, $(x \cup y)\downarrow$. Hence by theorem (145.2)★:

$$Class(x \cup y) \,\&\, Class(x) \,\&\, Class(y) \,\&\, UnionOf(x \cup y, x, y)$$

So by definition of *UnionOf* (350), the last conjunct implies:

$$\forall z(z \in x \cup y \equiv (z \in x \vee z \in y)) \qquad ⋈$$

(357.1) If we eliminate the restricted variables, then by GEN, we have to show:

$$Class(x) \to \exists y(Class(y) \,\&\, ComplementOf(y,x))$$

Assume *Class*(x), to find a witness to the consequent. By theorem (312.2), there is a property, say P, such that *ExtensionOf*(x,P). So consider the property $\overline{P}$, i.e., $[\lambda z\, \neg Pz]$, which we know exists by (39.2). Then by the Fundamental Theorem for Natural Classes and Logical Sets (318), we know:

$$\exists y(Class(y) \,\&\, \forall z(z \in y \equiv \overline{P}z))$$

Suppose a is an arbitrary such object, so that we know *Class*(a) and:

(ϑ) $\forall z(z \in a \equiv \overline{P}z)$

To see a that is our witness, it remains only to show $\forall z(z \in a \equiv z \notin x)$, by definition (356). Since z isn't free in any assumption, it suffices by GEN to show $z \in a \equiv z \notin x$. We do this as follows:

$$\begin{array}{lcll} z \in a & \equiv & \overline{P}z & \text{by } (\vartheta) \\ & \equiv & [\lambda z\, \neg Pz]z & \text{by definition } \overline{P}\text{, Rule } =_{df}\text{I, and Rule } =_{df}\text{E} \\ & \equiv & \neg Pz & \text{by Rule } \overrightarrow{\beta}\text{C (184.1.a) and (184.1.b)} \\ & \equiv & \neg z \in x & \text{by (317.1), } ExtensionOf(x,P) \\ & \equiv & z \notin x & \text{by convention for } \notin \end{array} \quad \bowtie$$

(357.2) By definition of the uniqueness quantifier, our theorem asserts:

$$\forall c \exists c'(ComplementOf(c',c) \,\&\, \forall c''(ComplementOf(c'',c) \to c''=c'))$$

Eliminating the restricted variables:

$$\begin{array}{l} \forall x(Class(x) \to \exists y(Class(y) \,\&\, ComplementOf(y,x) \,\&\, \\ \quad \forall z(Class(z) \,\&\, ComplementOf(z,x) \to z=y))) \end{array}$$

By GEN, we assume *Class*(x), and find a witness to the consequent. But we know, by (357.1), that since *Class*(x), $\exists y(Class(y) \,\&\, ComplementOf(y,x))$. Suppose a is such an object so that we know *Class*(a) and, by definition of *ComplementOf* (356):

(ϑ) $\forall w(w \in a \equiv w \notin x)$

To see that a is our witness, it remains to show $Class(z) \,\&\, ComplementOf(z,x) \to z=a$, by GEN. So assume *Class*(z) and *ComplementOf*(z,x). Then again by definition of *ComplementOf* (356):

(ξ) $\forall w(w \in z \equiv w \notin x)$

Hence by properties of the biconditional, (ϑ) and (ξ) imply $\forall w(w \in z \equiv w \in a)$. So by the principle of extensionality (343), $z=a$. $\bowtie$

(358)$\star$ If we eliminate the restricted variables and apply the Rule of Actualization to theorem (357.2), we know:

$$\mathcal{A}\forall x(Class(x) \to \exists!y(Class(y) \,\&\, ComplementOf(y,x)))$$

By the ⋆-theorem established at the beginning of footnote 231, it follows that:

$$\forall x(Class(x) \to \mathcal{A}\exists!y(Class(y) \,\&\, ComplementOf(y,x)))$$

By ∀E, it follows that:

$$Class(x) \to \mathcal{A}\exists!y(Class(y) \,\&\, ComplementOf(y,x))$$

Now assume $Class(x)$, for conditional proof. Then:

$$\mathcal{A}\exists!y(Class(y) \,\&\, ComplementOf(y,x))$$

So by theorem (176.2), it follows that $\imath y(Class(y) \& ComplementOf(y,x))\downarrow$. Hence, by conditional proof:

$$Class(x) \to \imath y(Class(y) \,\&\, ComplementOf(y,x))\downarrow$$

Since x isn't free in any undischarged assumptions, it follows by GEN that:

$$\forall x(Class(x) \to \imath y(Class(y) \,\&\, ComplementOf(y,x))\downarrow)$$

So by our conventions for restricted variables, $\forall c(\imath c' ComplementOf(c',c)\downarrow)$. ⋈

(360.1) By eliminating our restricted variables and GEN, we want to show:

$$(Class(x) \,\&\, Class(y)) \to \exists z(Class(z) \,\&\, IntersectionOf(z,x,y))$$

Assume $Class(x) \,\&\, Class(y)$, to find a witness to the consequent. Since x and y are classes, there are, by (312.2), properties P and Q such that $ExtensionOf(x,P)$ and $ExtensionOf(y,Q)$. Now consider the property $[\lambda z Pz \& Qz]$, which we know exists by (39.2). By the Fundamental Theorem for Natural Classes and Logical Sets (318), we know:

$$\exists x(Class(x) \,\&\, \forall w(w \in x \equiv [\lambda z\, Pz \,\&\, Qz]w))$$

Suppose a is an arbitrary such object, so that we know $Class(a)$ and $\forall w(w \in a \equiv [\lambda z\, Pz \,\&\, Qz]w)$. Then to see that a is our desired witness, it remains to show, By definition (359):

$$\forall w(w \in a \equiv (w \in x \,\&\, w \in y))$$

We leave this as an exercise. ⋈

(360.2) (Exercise)

(361)⋆ (Exercise)

(363)⋆ By our conventions for restricted variables and GEN, we want to show:

$$(Class(x) \,\&\, Class(y)) \to \forall z(z \in x \cap y \equiv (z \in x \,\&\, z \in y))$$

Assume $Class(x) \,\&\, Class(y)$. By our conventions for free restricted variables in definitions-by-= (339.1), definition of $\cap$ (362) abbreviates:

(ϑ) $x \cap y =_{df} \imath z(Class(z) \,\&\, Class(x) \,\&\, Class(y) \,\&\, IntersectionOf(z,x,y))$

Independently, since $Class(x) \,\&\, Class(y)$, it follows from (361)★ that:

(ξ) $\imath z(Class(z) \,\&\, IntersectionOf(z,x,y))\downarrow$

Similarly, by our conventions for free restricted variables in definitions-by-$\equiv$ (338.2), the definition of *IntersectionOf* (359) abbreviates:

$IntersectionOf(z,x,y) \equiv_{df} Class(z) \,\&\, Class(x) \,\&\, Class(y) \,\&\, \forall w(w \in z \equiv w \in x \vee w \in y)$

Since the definiens of $IntersectionOf(z,x,y)$ includes the conjuncts $Class(x)$ and $Class(y)$, it follows that (exercise):

$$\Box\forall z((Class(z) \,\&\, IntersectionOf(z,x,y)) \equiv \\ (Class(z) \,\&\, Class(x) \,\&\, Class(y) \,\&\, IntersectionOf(z,x,y)))$$

From this last result and (ξ), it follows by (149.3) that:

$$\imath z(Class(z) \,\&\, Class(x) \,\&\, Class(y) \,\&\, IntersectionOf(z,x,y))\downarrow$$

So by definition (ϑ) and the theories of definition and identity, $(x \cap y)\downarrow$. Hence by theorem (145.2)★:

$$Class(x \cap y) \,\&\, Class(x) \,\&\, Class(y) \,\&\, IntersectionOf(x \cap y, x, y)$$

So by definition of *IntersectionOf* (350), the last conjunct implies:

$$\forall z(z \in x \cup y \equiv (z \in x \,\&\, z \in y)) \qquad \bowtie$$

(364.1) When we eliminate the restricted variable, our theorem becomes:

$[\lambda y\, \varphi]\downarrow \to \exists z(Class(z) \,\&\, \forall y(y \in z \equiv \varphi))$, provided φ has no free occurrences of z

Assume $[\lambda y\, \varphi]\downarrow$, where z doesn't occur free in φ. We want to find a witness to the consequent. Now consider the following alphabetic variant of the Fundamental Theorem for Natural Classes and Logical Sets (318):

(ξ) $\forall F \exists z(Class(z) \,\&\, \forall y(y \in z \equiv Fy))$

Note that (a) $[\lambda y\, \varphi]$ exists; (b) z doesn't occur free in φ and so doesn't occur free in $[\lambda y\, \varphi]$; and (c) any occurrences of y in φ are bound by the λ and so y doesn't occur free in $[\lambda y\, \varphi]$. Hence, neither of the quantifiers $\exists z$ and $\forall y$ in (ξ) will capture any variables if we instantiate $[\lambda y\, \varphi]$ for $\forall F$ into (ξ). By doing so we obtain:

(ϑ) $\exists z(Class(z) \,\&\, \forall y(y \in z \equiv [\lambda y\, \varphi]y))$

But it also follows from the existence of $[\lambda y \varphi]$ that $[\lambda y \varphi] \equiv \varphi$ is a modally strict theorem, by β-Conversion. Hence by a Rule of Substitution, we may infer from (ϑ) that $\exists z(Class(z) \,\&\, \forall y(y \in z \equiv \varphi))$. ⋈

(364.2) (Exercise)

(364.3) When we eliminate the restricted variable, our theorem becomes:

> $[\lambda y\, \varphi]\!\downarrow \to \imath z(Class(z) \,\&\, \forall y(y \in z \equiv \varphi))\!\downarrow$, provided φ has no free occurrences of z

Assume $[\lambda y\, \varphi]\!\downarrow$ and that z doesn't occur free in φ. Then by the necessity of logical existence (106), $\Box[\lambda y\, \varphi]\!\downarrow$, and since necessity implies actuality (132):

(ϑ) $\mathcal{A}[\lambda y\, \varphi]\!\downarrow$

Now, independently, if we apply the Rule of Actualization to theorem (364.2), we obtain:

$$\mathcal{A}([\lambda y\, \varphi]\!\downarrow \to \exists! z(Class(z) \,\&\, \forall y(y \in z \equiv \varphi)))$$

This implies, by theorem (131):

(ξ) $\mathcal{A}[\lambda y\, \varphi]\!\downarrow \to \mathcal{A}\exists! z(Class(z) \,\&\, \forall y(y \in z \equiv \varphi^{)})$

It follows from (ξ) and (ϑ) that:

$$\mathcal{A}\exists! z(Class(z) \,\&\, \forall y(y \in z \equiv \varphi))$$

So by (176.2):

$$\imath z(Class(z) \,\&\, \forall y(y \in z \equiv \varphi))\!\downarrow$$

By our convention for restricted variables, $\imath c \forall y(y \in c \equiv \varphi)\!\downarrow$. ⋈

(368.1)★ Let $\varphi[y]$ be any y-predicable formula (365). Then if x doesn't occur free in φ, we may eliminate the restricted variables in the version of (364.3) formulated with our conventions in (365), to obtain:

(ϑ) $\imath x(Class(x) \,\&\, \forall y(y \in x \equiv \varphi[y]))\!\downarrow$

If x does occur free in φ, then pick some variable that is not free in φ, say z (without loss of generality), and use the alphabetic variant of (ϑ) with z replacing x as the bound variable in what follows. So by definition (366) and the Rule of Identity by Definition (120.1), we obtain:

(ζ) $\{y \mid \varphi[y]\} = \imath x(Class(x) \,\&\, \forall y(y \in x \equiv \varphi[y]))$

Since this implies $\{y \mid \varphi[y]\}\downarrow$, we may instantiate $\{y \mid \varphi[y]\}$ into (145.2)★ to produce the following instance:

$$(\xi)\ \ \{y \mid \varphi[y]\} = \imath x(Class(x) \,\&\, \forall y(y \in x \equiv \varphi[y])) \rightarrow \\ Class(\{y \mid \varphi[y]\}) \,\&\, \forall y(y \in \{y \mid \varphi[y]\} \equiv \varphi[y])$$

Hence, (ζ) and (ξ) imply:

$$\forall y(y \in \{y \mid \varphi[y]\} \equiv \varphi[y])$$

which suffices for the proof of our theorem by Rule $\forall$E. ⋈

(368.2)★ Let $\varphi[y]$ be any y-predicable formula (365) in which z is substitutable for y. By applying GEN to (368.1)★, we know:

$$(\vartheta)\ \ \forall y(y \in \{y \mid \varphi[y]\} \equiv \varphi[y])$$

Since z is substitutable for y in $\varphi[y]$, it is substitutable for y in $y \in \{y \mid \varphi[y]\} \equiv \varphi[y]$. Moreover, all free occurrences of y in $\varphi[y]$ become bound occurrences in $\{y \mid \varphi[y]\}$. So if we instantiate (ϑ) to z, we obtain:

$$z \in \{y \mid \varphi[y]\} \equiv \varphi[y]^z_y$$

by the definition of substitutions (14). ⋈

(369.1)[462] Let $\varphi[y]$ be any y-predicable formula (365), so that $[\lambda y\, \varphi[y]]\downarrow$. Now choose a variable, say x, that doesn't occur free in $\varphi[y]$. Then, by the discussion in Remark (367), we know that definitions (366) and (322) imply, respectively, by our Rule of Definition by Identity:

$$\{y \mid \varphi[y]\} = \imath x(Class(x) \,\&\, \forall y(y \in x \equiv \varphi[y]))$$

$$\epsilon[\lambda y\, \varphi[y]] = \imath x ExtensionOf(x, [\lambda y\, \varphi[y]])$$

Then to show $\{y \mid \varphi[y]\} = \epsilon[\lambda y\, \varphi[y]]$, it suffices to show:

$$\text{(A)}\ \ \imath x(Class(x) \,\&\, \forall y(y \in x \equiv \varphi[y])) = \imath x ExtensionOf(x, [\lambda y\, \varphi[y]])$$

To establish (A), we make use of theorem (149.3), which asserts:

$$\imath x\psi\downarrow \,\&\, \Box\forall x(\psi \equiv \chi) \rightarrow \imath x\psi = \imath x\chi$$

Now if we let ψ be $Class(x) \& \forall y(y \in x \equiv \varphi[y])$ and χ be $ExtensionOf(x, [\lambda y\, \varphi[y]])$, then to establish (A), we have to show:

---

[462] There is an easier, non-modally strict proof of this theorem that goes by way of the Principle of Extensionality. Since both $\{y \mid \varphi[y]\}$ and $\epsilon[\lambda y\, \varphi[y]]$ are classes, it suffices to show $z \in \{y \mid \varphi[y]\} \equiv z \in \epsilon[\lambda y\, \varphi[y]]$. But, for the $\rightarrow$ direction, if we assume $z \in \{y \mid \varphi[y]\}$, we'd have to use (368.2)★ to conclude $\varphi[y]^z_y$, and then go on to show $z \in \epsilon[\lambda y\, \varphi[y]]$. But the appeal to (368.2)★ undermines the modal strictness of the proof. The following proof is lengthy in part because it avoids (368.2)★.

(B) $\imath x(Class(x) \,\&\, \forall y(y \in x \equiv \varphi[y]))\downarrow$

(C) $\Box\forall x((Class(x) \,\&\, \forall y(y \in x \equiv \varphi[y])) \equiv ExtensionOf(x, [\lambda y\, \varphi[y]]))$

But (B) has already been established as theorem, by (364.3) and our conventions in (365). So it remains to show (C). By RN and GEN, it suffices to show:

(D) $(Class(x) \,\&\, \forall y(y \in x \equiv \varphi[y])) \equiv ExtensionOf(x, [\lambda y\, \varphi[y]])$

($\rightarrow$) Assume:

(E) $Class(x) \,\&\, \forall y(y \in x \equiv \varphi[y])$

Then from the first conjunct of (E) it follows by definition (312.2) that:

(F) $\exists G(ExtensionOf(x, G))$

Suppose P is such a property, so that we know:

(G) $ExtensionOf(x, P)$

Then by (317.1), (G) implies:

(H) $\forall y(y \in x \equiv Py)$

But the second conjunct of (E) and the fact that $[\lambda y\, \varphi[y]]\downarrow$ imply, by the relevant instance of β-Conversion (namely, $[\lambda y\, \varphi[y]]y \equiv \varphi[y]$) and the Rule of Substitution (160.2), that:

(I) $\forall y(y \in x \equiv [\lambda y\, \varphi[y]]y)$

So, from (H) and (I) we may conclude either of the following:

(J) $\forall y(Py \equiv [\lambda y\, \varphi[y]]y)$
$\forall y([\lambda y\, \varphi[y]]y \equiv Py)$

One final preliminary fact is that by (312.1), (G) implies:

(K) $A!x \,\&\, \forall F(xF \equiv \forall y(Fy \equiv Py))$

Now we want to show $ExtensionOf(x, [\lambda y\, \varphi[y]])$, i.e., by definition (312.1):

$$A!x \,\&\, [\lambda y\, \varphi[y]]\downarrow \,\&\, \forall F(xF \equiv \forall y(Fy \equiv [\lambda y\, \varphi[y]]y))$$

$A!x$ is the first conjunct of (K). $[\lambda y\, \varphi[y]]\downarrow$ is already known. So by GEN, it remains to show $xF \equiv \forall y(Fy \equiv [\lambda y\, \varphi[y]]y)$:

($\rightarrow$) Assume xF. From this and the second conjunct of (K), it follows that $\forall y(Fy \equiv Py)$. But this last result and (J) jointly imply $\forall y(Fy \equiv [\lambda y\, \varphi[y]]y)$.

($\leftarrow$) Assume $\forall y(Fy \equiv [\lambda y \varphi[y]]y)$. From this and (J) it follows that $\forall y(Fy \equiv Py)$. But then by (K), we may conclude xF.

(←) Assume $ExtensionOf(x, [\lambda y \varphi[y]])$. We have to show both (a) $Class(x)$ and (b) $\forall y(y \in x \equiv \varphi[y])$. (a) Our assumption implies $\exists G(ExtensionOf(x, G)$, and so by definition (312.2), $Class(x)$. (b) By (317.1), our assumption implies $\forall y(y \in x \equiv [\lambda y\, \varphi[y]]y)$. But since $[\lambda y\, \varphi[y]]\downarrow$, this implies $\forall y(y \in x \equiv \varphi[y])$ by β-Conversion and the Rule of Substitution. ⋈

(369.2) Gy is a y-predicable formula (365), since $[\lambda y\, Gy]\downarrow$ is an axiom and hence a theorem. So, as an instance of (369.1), we know that $\{y \mid Gy\} = \epsilon[\lambda y\, Gy]$. Now by η-Conversion (48.3), $[\lambda y\, Gy] = G$. Hence, by Rule =E, $\{y \mid Gy\} = \epsilon G$. ⋈

(370.1) By hypothesis, $\varphi[y]$ is y-predicable formula, and so it is a theorem that the property $[\lambda y\, \varphi[y]]$ exists. If we eliminate the restricted variables from our theorem, then we want to show, by GEN:

$$Class(x) \to \exists z(Class(z) \,\&\, \forall y(y \in z \equiv y \in x \,\&\, \varphi[y])),$$
provided $\varphi[y]$ has no free occurrences of z

Assume $Class(x)$, to find a witness to the consequent. Then there exists a property, say P, such that $ExtentionOf(x, P)$, by (312.2). So by (317.1), we know the following:

(ϑ) $\forall y(y \in x \equiv Py)$

Now pick some variable, say w, that doesn't occur free in $\varphi[y]$ and consider the property $[\lambda w\, Pw \,\&\, [\lambda y\, \varphi[y]]w]$. To see that this property exists, note that by (39.2), $[\lambda w\, Pw \,\&\, Fw]\downarrow$, for every property F. So, since $[\lambda y\, \varphi[y]]$ exists, we can instantiate it for F to obtain $[\lambda w\, Pw \,\&\, [\lambda y\, \varphi[y]]w]\downarrow$. Then by the Fundamental Theorem for Natural Classes and Logical Sets (318), we therefore know:

$$\exists z(Class(z) \,\&\, \forall y(y \in z \equiv [\lambda w\, Pw \,\&\, [\lambda y\, \varphi[y]]w]y)$$

Suppose a is such an object, so that we know $Class(a)$ and:

(A) $\forall y(y \in a \equiv [\lambda w\, Pw \,\&\, [\lambda y\, \varphi[y]]w]y)$

To confirm that a is our desired witness, it remains only to show:

(B) $\forall y(y \in a \equiv y \in x \,\&\, \varphi[y])$

Since $[\lambda w\, Pw \,\&\, [\lambda y\, \varphi[y]]w]\downarrow$, β-Conversion yields the modally strict fact that:

(C) $[\lambda w\, Pw \,\&\, [\lambda y\, \varphi[y]]w]y \equiv Py \,\&\, [\lambda y\, \varphi[y]]y$

Then by a Rule of Substitution, (A) and (C) imply:

(D) $\forall y(y \in a \equiv Py \,\&\, [\lambda y\, \varphi[y]]y)$

But $[\lambda y\, \varphi[y]]$ also exists, and so β-Conversion also yields the modally strict fact:

(E) $[\lambda y\, \varphi[y]]y \equiv \varphi[y]$

So by a Rule of Substitution, (D) and (E) imply:

(F) $\forall y(y \in a \equiv Py \,\&\, \varphi[y])$

But now to show (B), it suffices to show:

(G) $\forall y(Py \,\&\, \varphi[y] \equiv y \in x \,\&\, \varphi[y])$

since (F) and (G) imply (B) by the properties of the quantified biconditional (99.10). So by GEN, we need to show:

$$Py \,\&\, \varphi[y] \equiv y \in x \,\&\, \varphi[y]$$

But this follows easily from a consequence of (ϑ), namely $Py \equiv y \in x$, by (88.4.e). ⋈

(370.2) (Exercise)

(371)★ (Exercise)

(373)★ Let $\varphi[y]$ be any y-predicable formula in which z is substitutable for y. If we eliminate the restricted variable from our theorem, then we have to show, by GEN:

$$Class(x) \to \forall z(z \in \{y \mid y \in x \,\&\, \varphi[y]\} \equiv (z \in x \,\&\, \varphi[y]^z_y))$$

Assume $Class(x)$, to find a witness to the consequent. Since x is a class, we may instantiate it into (371)★, and if we eliminate the restricted variable in the result by choosing some variable, say w, that doesn't occur free in $\varphi[y]$, we know:

$$\imath w(Class(w) \,\&\, \forall y(y \in w \equiv y \in x \,\&\, \varphi[y]))\downarrow$$

Hence, by definition (372) and the Rule of Identity by Definition (120.1):

$$\{y \mid y \in x \,\&\, \varphi[y]\} = \imath w(Class(w) \,\&\, \forall y(y \in w \equiv y \in x \,\&\, \varphi[y]))$$

Since this implies that the class abstract is significant, it follows by (145)★ that:

(ϑ) $Class(\{y \mid y \in x \,\&\, \varphi\}) \,\&\, \forall y(y \in \{y \mid y \in x \,\&\, \varphi[y]\} \equiv (y \in x \,\&\, \varphi[y]))$

By hypothesis, z is substitutable for y in $\varphi[y]$. Hence it is substitutable for y in the matrix of the second conjunct of (ϑ). So as an alphabetic variant of the second conjunct of (ϑ), we have:

$$\forall z(z \in \{y \mid y \in x \,\&\, \varphi[y]\} \equiv (z \in x \,\&\, \varphi[y]^z_y))$$ ⋈

(374)★ By hypothesis, $\varphi[y]$ is a y-predicable formula. Assume $Class(x)$. Then we have to show:

$$\{y \mid y \in x \,\&\, \varphi[y]\} = x \cap \{y \mid \varphi[y]\}$$

By the Principle of Extensionality (343), it suffices to show:

$$\forall z(z \in \{y \mid y \in x \,\&\, \varphi[y]\} \equiv z \in x \cap \{y \mid \varphi[y]\})$$

By GEN, it suffices to show:

$$z \in \{y \mid y \in x \,\&\, \varphi[y]\} \equiv z \in x \cap \{y \mid \varphi[y]\}$$

($\rightarrow$) Assume $z \in \{y \mid y \in x \,\&\, \varphi[y]\}$. Then by (373)★, it follows that $z \in x \,\&\, \varphi[y]^z_y$. But by (368.2)★, the second conjunct is equivalent to $z \in \{y|\varphi[y]\}$. Hence $z \in x \,\&$ $z \in \{y \mid \varphi[y]\}$. So by (363)★, $z \in x \cap \{y \mid \varphi[y]\}$. ($\leftarrow$) By reversing the reasoning. ⋈

(375) Eliminating the restricted variable, our theorem asserts:

$$\forall R \forall x(\mathit{Class}(x) \rightarrow \exists z(\mathit{Class}(z) \,\&\, \forall y(y \in z \equiv \exists w(w \in x \,\&\, Rwy))))$$

By GEN, we assume $\mathit{Class}(x)$, to find a witness. Then $\mathit{ExtensionOf}(x,P)$, for some P. Hence by (317.1), it follows that:

(ϑ) $\forall y(y \in x \equiv Py)$

Now consider the property $[\lambda x \exists w(Pw \,\&\, Rwx)]$, which we know exists by (39.2). By the Fundamental Theorem for Natural Classes:

$$\exists x(\mathit{Class}(x) \,\&\, \forall y(y \in x \equiv [\lambda x \exists w(Pw \,\&\, Rwx)]y))$$

Suppose a is such an object, so that we know $\mathit{Class}(a)$ and:

(A) $\forall y(y \in a \equiv [\lambda x \exists w(Pw \,\&\, Rwx)]y)$

To see a is our desired witness, it remains only to show:

(B) $\forall y(y \in a \equiv \exists w(w \in x \,\&\, Rwy))$

Since $[\lambda x \exists w(Pw \,\&\, Rwx)]\downarrow$, β-Conversion yields the modally strict fact that:

(C) $[\lambda x \exists w(Pw \,\&\, Rwx)]y \equiv \exists w(Pw \,\&\, Rwy)$

Then by a Rule of Substitution, (A) and (C) imply:

(D) $\forall y(y \in a \equiv \exists w(Pw \,\&\, Rwy))$

But to show (B), it suffices to show:

(E) $\forall y(\exists w(Pw \,\&\, Rwy) \equiv \exists w(w \in x \,\&\, Rwy))$

since (D) and (E) imply (B) by the properties of the quantified biconditional (99.10). So by GEN, we need to show:

$$\exists w(Pw \,\&\, Rwy) \equiv \exists w(w \in x \,\&\, Rwy)$$

($\rightarrow$) Assume $\exists w(Pw \,\&\, Rwy)$. Suppose b such an individual, so that we know $Pb \,\&\, Rby$. Instantiating b to (ϑ), we know $b \in x \equiv Pb$, i.e., $Pb \equiv b \in x$. Hence, by (88.4.e), it follows that $b \in x \,\&\, Rby$. So $\exists w(w \in x \,\&\, Rwy)$.

($\leftarrow$) Assume $\exists w(w \in x \,\&\, Rwy)$. Suppose b is such an individual, so that we know $b \in x \,\&\, Rby$. Again, instantiating to (ϑ), we know $b \in x \equiv Pb$. Hence by (88.4.e), it follows that $Pb \,\&\, Rby$. Hence, $\exists w(Pw \,\&\, Rwy)$. ⋈

(376) For reductio, assume:

$$\forall x \exists c \forall y(y \in c \equiv y = x),$$

i.e.,

$$\forall x \exists z(Class(z) \,\&\, \forall y(y \in z \equiv y = x)$$

Instantiate this to an arbitrary x, so that we know:

$$\exists z(Class(z) \,\&\, \forall y(y \in z \equiv y = x)$$

Now suppose a is such an object, so that we know:

$$(\vartheta) \quad Class(a) \,\&\, \forall y(y \in a \equiv y = x)$$

$Class(a)$ implies, by definition (312.2), that $\exists G(ExtensionOf(a, G))$. Suppose then that that $ExtensionOf(a, P)$. Then by (317.1), it follows that:

$$(\xi) \quad \forall y(y \in a \equiv Py)$$

Then the 2nd conjunct of (ϑ) and (ξ) imply:

$$\forall y(Py \equiv y = x)$$

By $\exists$I, it follows that $\exists F \forall y(Fy \equiv y = x)$. And since x was arbitrary and doesn't occur free in any assumption, it follows by GEN that $\forall x \exists F \forall y(Fy \equiv y = x)$, which contradicts (192.5). ⋈

(377.1) We have to show:

$$\exists c \forall y(y \in c \equiv D!y \,\&\, y = x)$$

But by (364.1), it suffices to show $[\lambda y\, D!y \,\&\, y = x]\downarrow$. But this is an instance of (273.13). ⋈

(377.2) (Exercise)

(377.3) By (377.2), we know:

$$\exists! z(Class(z) \,\&\, \forall y(y \in z \equiv D!y \,\&\, y = x))$$

So by the Rule of Actualization:

$$\mathcal{A}\exists!z(Class(z) \,\&\, \forall y(y \in z \equiv D!y \,\&\, y = x))$$

Hence by (176.2):

$$\imath z(Class(z) \,\&\, \forall y(y \in z \equiv D!y \,\&\, y = x))\downarrow$$

So by our conventions for restricted variables, $\imath c\forall y(y \in c \equiv D!y \,\&\, y = x))\downarrow$. ⋈

(377.4) (Exercise)

(379.1) A non-modally strict proof is relatively easy, but the following, modally strict proof is more involved. Assume $A!z$ and $\neg D!z$. We want to show $\{z\}_D = \varnothing$. We know both objects exist: $\{z\}_D$ exists by (378) and (377.4), and $\varnothing$ exists by (346) and (345.3). So by definitions (378) and (346), respectively, we have to show:

$$\{y \,|\, D!y \,\&\, y = z\} = \imath x(Class(x) \,\&\, Empty(x))$$

By definition (366), we have to show:

$$\imath x(Class(x) \,\&\, \forall y(y \in x \equiv D!y \,\&\, y = z)) = \imath x(Class(x) \,\&\, Empty(x))$$

To prove this by modally strict means, our proof strategy is to derive the above from an appropriate instance of (149.2), which says that if $\imath x\varphi\downarrow \,\&\, \mathcal{A}\forall x(\varphi \equiv \psi) \rightarrow \imath x\varphi = \imath x\psi$. So if we set:

$$\varphi = Class(x) \,\&\, \forall y(y \in x \equiv D!y \,\&\, y = z)$$

$$\psi = Class(x) \,\&\, Empty(x)$$

then since $\imath x\varphi\downarrow$ is already known, it remains only to show that we can derive:

$$\mathcal{A}\forall x(\varphi \equiv \psi)$$

from our assumptions that $A!z$ and $\neg D!z$. By Rule CP, it suffices to show:

$$A!z, \neg D!z \vdash \mathcal{A}\forall x(\varphi \equiv \psi)$$

Note that the following is an instance of the Rule of Actualization (RA):

If $A!z, \neg D!z, \varphi \vdash \psi$, then $\mathcal{A}A!z, \mathcal{A}\neg D!z, \mathcal{A}\varphi \vdash \mathcal{A}\psi$

Moreover, as an instance of Rule CP (75), we know:

If $\mathcal{A}A!z, \mathcal{A}\neg D!z, \mathcal{A}\varphi \vdash \mathcal{A}\psi$, then $\mathcal{A}A!z, \mathcal{A}\neg D!z \vdash \mathcal{A}\varphi \rightarrow \mathcal{A}\psi$

So, combining these two facts, the following is a valid form of reasoning:

(A) If $A!z, \neg D!z, \varphi \vdash \psi$, then $\mathcal{A}A!z, \mathcal{A}\neg D!z \vdash \mathcal{A}\varphi \rightarrow \mathcal{A}\psi$

And by analogous reasoning:

(A′) If $A!z, \neg D!z, \psi \vdash \varphi$, then $\mathcal{A}A!z, \mathcal{A}\neg D!z \vdash \mathcal{A}\psi \to \mathcal{A}\varphi$

Our strategy now is as follows:

(B) Show $A!z, \neg D!z, \varphi \vdash \psi$

(B′) Show $A!z, \neg D!z, \psi \vdash \varphi$.

(C) Infer $\mathcal{A}A!z, \mathcal{A}\neg D!z \vdash \mathcal{A}\varphi \equiv \mathcal{A}\psi$, from (B) and (A), (B′) and (A′), by &I and the definition of $\equiv$.

(D) Infer $\mathcal{A}A!z, \mathcal{A}\neg D!z \vdash \mathcal{A}(\varphi \equiv \psi)$ from (C), by the right-to-left direction of (139.5).

(E) Infer $A!z, \neg D!z \vdash \mathcal{A}(\varphi \equiv \psi)$, from (D) and the facts that $A!z \vdash \mathcal{A}A!z$ and $D!z \vdash \mathcal{A}D!z$; these latter facts are consequences of left-to-right directions of (180.8) and (273.12), and (63.10).

(F) Infer $A!z, \neg D!z \vdash \forall x \mathcal{A}(\varphi \equiv \psi)$, by GEN, from (E), since x isn't free in the assumption $A!z$.

(G) Conclude $A!z, \neg D!z \vdash \mathcal{A}\forall x(\varphi \equiv \psi)$, from (F) and axiom (44.3).

Since steps (C) – (G) are straightforward if not explicit, it remains only to establish (B) and (B′):

(B) Assume $A!z$, $\neg D!z$, and φ, i.e., $Class(x) \,\&\, \forall y(y \in x \equiv D!y \,\&\, y = z)$. To show ψ, it suffices to show $Empty(x)$, and by definition (344), it suffices to show $\neg\exists y(y \in x)$. Assume, for reductio, $\exists y(y \in x)$. Suppose a is such an object, so that we know $a \in x$. Then $D!a \,\&\, a = z$. But this last fact implies $D!z$. Contradiction.

(B′) Assume $A!z$, $\neg D!z$, and ψ, i.e., $Class(x) \,\&\, Empty(x)$. To show φ, it suffices, by GEN, to show $y \in x \equiv D!y \,\&\, y = z$. Since x is empty, it follows by definition (344) that $\neg\exists y(y \in x)$, i.e., $\forall y \neg(y \in x)$. ($\to$) Hence $\neg(y \in x)$ and so $y \in x \to D!y \,\&\, y = z$, by the failure of the antecedent. ($\leftarrow$) Since $\neg D!z$, it follows that $\neg(D!y \,\&\, y = z)$ on pain of contradiction. So by failure of the antecedent, $(D!y \,\&\, y = z) \to y \in x$. ⋈

(379.2) Assume, for reductio, $\forall x \forall z(x \neq z \to \{x\}_D \neq \{z\}_D)$. Now by (269), we know:

$$\exists x \exists y(A!x \,\&\, A!y \,\&\, x \neq y \,\&\, \forall F(Fx \equiv Fy))$$

Let a and b be such objects, so that we know:

$$A!a \,\&\, A!b \,\&\, a \neq b \,\&\, \forall F(Fa \equiv Fb)$$

Since $a \neq b$, it follows by our reductio assumption that $\{a\}_D \neq \{b\}_D$. But since both $\forall F(Fa \equiv Fb)$ and $a \neq b$, we also know $\neg(\forall F(Fa \equiv Fb) \rightarrow a \neq b)$. So by (273.7), it follows that both $\neg D!a$ and $\neg D!b$. By (379.1), these last two facts imply, respectively, that $\{a\}_D = \varnothing$ and $\{b\}_D = \varnothing$. Hence, $\{a\}_D = \{b\}_D$. Contradiction. ⋈

(379.3) Assume $D!z$. For reductio, suppose $\{z\}_D = \varnothing$. We know independently, by definitions (378), (366) and (346), that:

(ϑ) $\{z\}_D = \imath x(Class(x) \,\&\, \forall y(y \in x \equiv y =_D z))$

(ξ) $\varnothing = \imath x(Class(x) \,\&\, Empty(x))$

From (ϑ), it follows by theorem (152.3) that:

$$\mathcal{A}(Class(\{z\}_D) \,\&\, \forall y(y \in \{z\}_D \equiv y =_D z))$$

So given our reductio hypothesis:

$$\mathcal{A}(Class(\varnothing) \,\&\, \forall y(y \in \varnothing \equiv y =_D z))$$

If we distribute the actuality operator over the conjunction and commute $\mathcal{A}$ and $\forall y$ on the second conjunct of the result, it follows *a fortiori* that:

$$\forall y \mathcal{A}(y \in \varnothing \equiv y =_D z)$$

Hence:

$$\mathcal{A}(z \in \varnothing \equiv z =_D z)$$

But it follows from our initial assumption that $z =_D z$ (273.30), and so $\mathcal{A}z =_D z$, by (273.24). So by (139.5):

(ζ) $\mathcal{A}z \in \varnothing$

But this can't be. From (ξ), it follows that $\mathcal{A}(Class(\varnothing) \,\&\, Empty(\varnothing))$, by theorem (152.3). By distributing the actuality operator, it follows that $\mathcal{A}Empty(\varnothing)$. By definition (344) and the Rule of Substitution (160.2), this implies $\mathcal{A}\neg\exists y(y \in \varnothing)$, i.e., $\mathcal{A}\forall y\neg(y \in \varnothing)$. So by commutativity, $\forall y\mathcal{A}\neg(y \in \varnothing)$. Hence $\mathcal{A}\neg(z \in \varnothing)$, which by (44.1) implies $\neg\mathcal{A}z \in \varnothing$. This contradicts ($\zeta$). ⋈

(379.4) Assume $D!x$, $D!z$, and $x \neq z$. Note that if we eliminate the infix notation for $\neq$ (24), then $x \neq z$ implies $\neg(x =_D z)$, by (273.19). Now for reductio, suppose $\{x\}_D = \{z\}_D$. By applying definitions, we know:

(ϑ) $\{x\}_D = \imath w(Class(w) \,\&\, \forall y(y \in w \equiv y =_D x))$

(ξ) $\{z\}_D = \imath w'(Class(w') \,\&\, \forall y(y \in w' \equiv y =_D z))$

From (ξ), it follows by theorem (152.3) that:

$$\mathcal{A}(Class(\{z\}_D) \,\&\, \forall y(y \in \{z\}_D \equiv y =_D z))$$

If we distribute the actuality operator over the conjunction and commute $\mathcal{A}$ and $\forall y$ in the second conjunct of the result, it follows *a fortiori* that:

$$\forall y \mathcal{A}(y \in \{z\}_D \equiv y =_D z)$$

Instantiating to z:

$$\mathcal{A}(z \in \{z\}_D \equiv z =_D z)$$

But from $D!z$, it follows that $z =_D z$ (273.30), and so by (273.24), $\mathcal{A}z =_E z$. From this and our last displayed result, it follows by (139.5) that :

(ζ) $\mathcal{A}z \in \{z\}_D$

Put (ζ) aside for the moment and note that (ϑ) implies, by theorem (152.3), that:

$$\mathcal{A}(Class(\{x\}_D) \,\&\, \forall y(y \in \{x\}_D \equiv y =_D x))$$

Given our reductio hypothesis, this implies:

$$\mathcal{A}(Class(\{z\}_D) \,\&\, \forall y(y \in \{z\}_D \equiv y =_D x))$$

By now familiar reasoning, we can infer from this that:

$$\mathcal{A}(z \in \{z\}_D \equiv z =_D x)$$

From this and (ζ), it follows that $\mathcal{A}z =_D x$. But this implies, by (273.24), that $z =_D x$, which by (273.31), implies $x =_D z$. Contradiction. $\bowtie$

(380)★ Assume $D!x$. Clearly, $\{x\}_D\downarrow$ (exercise). Now, by GEN, we have to show $y \in \{x\}_D \equiv y = x$. We may therefore reason as follows:

$y \in \{x\}_D$	$\equiv$	$y \in \{y \mid D!y \,\&\, y = x\}$	by definition of $\{x\}_D$ (378)
	$\equiv$	$D!y \,\&\, y = x$	by (368.1)★
	$\equiv$	$y = x$	($\rightarrow$) by &E; ($\leftarrow$) by $D!x$ and Rule =E $\bowtie$

(381.1) – **(381.2)** (Exercises)

(382.1) If we eliminate the restricted variables, then we have to show:

$$\forall z(Class(z) \rightarrow \exists w(Class(w) \,\&\, \forall y(y \in w \equiv y \in z \vee y =_D x)))$$

By GEN, assume $Class(z)$, to find a witness to the consequent. Then there is some property, say P, such that $ExtensionOf(z, P)$.[463] Now consider the property $[\lambda y\, Py \vee y =_D x]$. This property clearly exists. Then by the Fundamental Theorem:

[463] It is tempting, at this point, to consider the class $\{y \mid Py \vee y =_P x\}$ and then show it is a witness that proves the theorem. But, then, we would have to argue by way of (368.1)★, which provides us with the fact:

(ϑ) $y \in \{y \mid Py \vee y =_D x\} \equiv (Py \vee y =_D x)$

Though we could subsequently prove that $y \in \{y \mid Py \vee y =_D x\} \equiv y \in z \vee y =_D x$, which by GEN and $\exists$I, yields the desired conclusion. But the resulting proof wouldn't be modally strict.

$$\exists w(Class(w) \,\&\, \forall y(y \in w \equiv [\lambda y\, Py \vee y =_D x]y))$$

Suppose a is an arbitrary such class. By now familiar reasoning, to show a is our witness, it remains to show $y \in a \equiv y \in z \vee y =_D x$. (Exercise) ⋈

(**382**.2) (Exercise)

(**386**.1) – (**386**.4) (Exercises)

(**387**.1) – (**387**.2) (Exercises)

(**389**.1)★ – (**389**.2)★ (Exercises)

(**391**) By hypothesis, $\|$ is an equivalence relation with respect to lines. Then if we eliminate the restricted variables, we have to show:

$$\forall x \forall x'((Lx \,\&\, Lx') \rightarrow (\forall z([\lambda y\, Ly \,\&\, y \| x]z \equiv [\lambda y\, Ly \,\&\, y \| x']z) \equiv x \| x')$$

By GEN, assume Lx and Lx'. ($\rightarrow$) Assume:

$$\forall z([\lambda y\, Ly \,\&\, y \| x]z \equiv [\lambda y\, Ly \,\&\, y \| x']z)$$

Then, instantiating to x:

$$[\lambda y\, Ly \,\&\, y \| x]x \equiv [\lambda y\, Ly \,\&\, y \| x']x$$

Since both $[\lambda y\, Ly \,\&\, y \| x]$ and $[\lambda y\, Ly \,\&\, y \| x']$ exist, we can reduce both the left and right conditions by strengthened β-Conversion and a Rule of Substitution:

$$(Lx \,\&\, x \| x) \equiv (Lx \,\&\, x \| x')$$

But this is equivalent to (88.8.e):

$$Lx \rightarrow (x \| x \equiv x \| x')$$

Since Lx by assumption, it follows that $x \| x \equiv x \| x'$. And since $\|$ is, by hypothesis, an equivalence relation, we know $x \| x$. Hence $x \| x'$.
($\leftarrow$) Assume $x \| x'$. We want to show:

$$\forall z([\lambda y\, Ly \,\&\, y \| x]z \equiv [\lambda y\, Ly \,\&\, y \| x']z)$$

By GEN, it suffices to show:

$$[\lambda y\, Ly \,\&\, y \| x]z \equiv [\lambda y\, Ly \,\&\, y \| x']z$$

We prove both directions:

($\rightarrow$) Assume $[\lambda y\, Ly \,\&\, y \| x]z$. Then Rule $\overrightarrow{\beta}$C yields both Lz and $z \| x$. From the latter, our assumption that $x \| x'$, and the transitivity of $\|$, it follows that $z \| x'$. Since $[\lambda y\, Ly \,\&\, y \| x']$ exists and Lz, Rule $\overleftarrow{\beta}$C yields $[\lambda y\, Ly \,\&\, y \| x']z$.

($\leftarrow$) Assume $[\lambda y\, Ly \,\&\, y \,\|\, x']z$. So by Rule $\overrightarrow{\beta}$C, we know both Lz and $z \,\|\, x'$. Independently, by the symmetry of $\|$, our assumption $x \,\|\, x'$ implies that $x' \,\|\, x$. Hence, by the transitivity of $\|$, $z \,\|\, x$. So from the existence of the terms in question and the fact that Lz, it follows that $[\lambda y\, Ly \,\&\, y \,\|\, x]z$, by Rule $\overleftarrow{\beta}$C. ⋈

(393.1) Eliminating the restricted variables, we have to show:

$$\forall y(Ly \rightarrow \exists x DirectionOf(x,y))$$

So by GEN, assume Ly. By definition (392), we have to find a witness to:

$$\exists x(Ly \,\&\, ExtensionOf(x,[\lambda z\, Lz \,\&\, z \| y])$$

Clearly, $[\lambda z\, Lz \,\&\, z \| y]$ exists. So $\exists x ExtensionOf(x,[\lambda z\, Lz \,\&\, z \| y])$, by (315.1). Let a be such an object, so that we have $ExtensionOf(a,[\lambda z\, Lz \,\&\, z \| y])$. Hence:

$$Ly \,\&\, ExtensionOf(a,[\lambda z\, Lz \,\&\, z \| y])$$

Then a is our witness. ⋈

(393.2) (Exercise)

(394) (Exercise)

(396)★ (Exercise)

(397)★ If we eliminate the restricted variables and apply the Rule of Actualization to theorem (393.2), we know $\mathcal{A}\forall y(Ly \rightarrow \exists! x DirectionOf(x,y))$. By the 2nd of the two ★-theorems established at the beginning of footnote 231, it follows that $\forall y(Ly \rightarrow \mathcal{A}\exists! x DirectionOf(x,y))$. So by $\forall$E, $Ly \rightarrow \mathcal{A}\exists! x DirectionOf(x,y)$. Now assume Ly, for conditional proof. Then $\mathcal{A}\exists! x DirectionOf(x,y)$. So by (176.2), $\imath x DirectionOf(x,y)\downarrow$. Hence, $Ly \rightarrow \imath x DirectionOf(x,y)\downarrow$, by conditional proof. Since y isn't free in any assumption, it follows by GEN that:

$$\forall y(Ly \rightarrow \imath x DirectionOf(x,y)\downarrow)$$

Employing our conventions for restricted variables: $\forall u(\imath x DirectionOf(x,u)\downarrow)$. ⋈

(399)★ Eliminating the restricted variables, we want to show:

$$\forall y \forall z((Ly \,\&\, Lz) \rightarrow (\vec{y} = \vec{z} \equiv y \| z))$$

By GEN, assume Lx and Ly. By definition (398), we know both:

$$\vec{y} = \imath x(Ly \,\&\, DirectionOf(x,y))$$

$$\vec{z} = \imath x(Lz \,\&\, DirectionOf(x,z))$$

Since $\vec{y}$ and $\vec{z}$ both exist, it follows by (145.2)★, respectively, that:

$$Ly \,\&\, DirectionOf(\vec{y}, y)$$

$$Ly \,\&\, DirectionOf(\vec{z}, z)$$

It follows from the second conjuncts of the last two results, respectively, by definition (392), that:

$$ExtensionOf(\vec{y}, [\lambda w\, Lw \,\&\, w \| y])$$

$$ExtensionOf(\vec{z}, [\lambda w\, Lw \,\&\, w \| z])$$

Hence, by (327.3)★, respectively:

$$(\vartheta)\quad \vec{y} = \epsilon[\lambda w\, Lw \,\&\, w \| y]$$

$$(\xi)\quad \vec{z} = \epsilon[\lambda w\, Lw \,\&\, w \| z]$$

So we may now reason as follows:

$$\begin{array}{llll} \vec{y} = \vec{z} & \equiv & \epsilon[\lambda w\, Lw \,\&\, w \| y] = \epsilon[\lambda w\, Lw \,\&\, w \| z] & \text{by } (\vartheta), (\xi), \text{ Rule =E } (\times 2) \\ & \equiv & \forall x([\lambda w\, Lw \,\&\, w \| y]x \equiv [\lambda w\, Lw \,\&\, w \| z]x) & \text{by Basic Law V (328)}\bigstar \\ & \equiv & y \| z & \text{by (391)} \quad \bowtie \end{array}$$

(401) (Exercise)

(403.1) – **(403.2)** (Exercises)

(404) (Exercise)

(406)★ (Exercise)

(407)★ (Exercise)

(409)★ (Exercise)

(411) (Exercise)

(412.1) By ∃I, from (411) and the fact $[\lambda xy\, \forall F(Fx \equiv Fy)]\downarrow$. ⋈.

(412.2) Let Π be any binary relation term. Assume *Equivalence*(Π). Then it follows from definition (410) that Π is reflexive, symmetric, and transitive. By reflexivity, $\forall x \Pi xx$. Then Πxx, by ∀E. Hence $\Pi\downarrow$, by axiom (39.5.a). ⋈

(413.1) If we eliminate the restricted variable, then by GEN, we have to show:

$$Equivalence(F) \rightarrow (\forall w([\lambda z\, Fzx]w \equiv [\lambda z\, Fzy]w) \equiv Fxy)$$

So assume $Equivalence(F)$. ($\rightarrow$) Assume $\forall w([\lambda z\, Fzx]w \equiv [\lambda z\, Fzy]w)$. Instantiating to x, it follows that $[\lambda z\, Fzx]x \equiv [\lambda z\, Fzy]x$. Since both $[\lambda z\, Fzx]$ and $[\lambda z\, Fzx]$ exist, we can reduce both the left and right conditions by β-Conversion. Hence we may conclude $Fxx \equiv Fxy$. But F is, by hypothesis, an equivalence relation, and so Fxx. Hence, Fxy. ($\leftarrow$) Exercise. ⋈

($\leftarrow$) Let Fxy be our global assumption. We want to show $\forall w([\lambda z\, Fzx]w \equiv [\lambda z\, Fzy]w)$. By GEN, it suffices to show $[\lambda z\, Fzx]w \equiv [\lambda z\, Fzy]w$:

> ($\rightarrow$) Assume $[\lambda z\, Fzx]w$. So by Rule $\overrightarrow{\beta}$C, Fwx. Since F is a transitive relation, this last result and our global assumption imply Fwy. Since $[\lambda z\, Fzy]$ and w exist, it follows by Rule $\overleftarrow{\beta}$C that $[\lambda z\, Fzy]w$.
>
> ($\leftarrow$) Assume $[\lambda z\, Fzy]w$. So by Rule $\overrightarrow{\beta}$C, Fwy. Independently, by the symmetry of F, our global assumption Fxy implies that Fyx. Hence Fwx, by the transitivity of F. So by Rule $\overleftarrow{\beta}$C, $[\lambda z\, Fzx]w$. ⋈

(413.2) – **(413.3)** (Exercises)

(415) (Exercise)

(416.1) – **(416.2)** (Exercises)

(417)★ (Exercise)

(419)★ If we eliminate the restricted variable, then by GEN, we have to show:

$$Equivalence(F) \rightarrow (\widehat{x}_F = \widehat{y}_F \equiv Fxy)$$

We reason by analogy with (399)★. By definition (418) and our conventions for restricted variables, we know both:

$$\widehat{x}_F = \imath w(Equivalence(F)\ \&\ F\text{-}AbstractionOf(w,x))$$

$$\widehat{y}_F = \imath w(Equivalence(F)\ \&\ F\text{-}AbstractionOf(w,y))$$

Since $\widehat{x}_F$ and $\widehat{y}_F$ both exist, it follows by (145.2)★, respectively, that:

$$Equivalence(F)\ \&\ F\text{-}AbstractionOf(\widehat{x}_F,x)$$

$$Equivalence(F)\ \&\ F\text{-}AbstractionOf(\widehat{y}_F,y)$$

But since both $[\lambda z\, Fzx]$ and $[\lambda z\, Fzy]$ exist, the second conjuncts of each of the above respectively imply, by definition (414), that:

$$ExtensionOf(\widehat{x}_F,[\lambda z\, Fzx])$$

$$ExtensionOf(\widehat{y}_F,[\lambda z\, Fzy])$$

Hence, by (327.3)★, respectively:

(ϑ) $\widehat{x}_F = \epsilon[\lambda z\, Fzx]$

(ξ) $\widehat{y}_F = \epsilon[\lambda z\, Fzy]$

So we may now reason as follows:

$$\begin{array}{llll} \widehat{x}_F = \widehat{y}_F & \equiv & \epsilon[\lambda z\, Fzx] = \epsilon[\lambda z\, Fzy] & \text{by } (\vartheta), (\xi), \text{Rule} =\text{E} \\ & \equiv & \forall w([\lambda z\, Fzx]w \equiv [\lambda z\, Fzy]w) & \text{by Basic Law V (328)}\star \\ & \equiv & Fxy & \text{by (413.1)} \qquad \bowtie \end{array}$$

(422.1) – **(422.4)** (Exercises)

(424) (Exercise) [Hint: The proof is analogous to the proofs of theorems (296.1) and (324.1).]

(425) By GEN, it suffices to show: $F = G \to \Box F = G$. But this is just an instance of the necessity of identity (125.1). (This argument reprises the one in Remark (262), in the discussion of Example (b).) ⋈

(426.1) By (424), we know $\boldsymbol{a}_G$ is the abstract object that encodes just G:

$$\boldsymbol{a}_G = \imath x(A!x \,\&\, \forall F(xF \equiv F = G))$$

Hence $\boldsymbol{a}_G\!\downarrow$. Moreover, we know by (425) that $F = G$ is a rigid condition on properties. So by (261.2):

$$A!\boldsymbol{a}_G \,\&\, \forall F(\boldsymbol{a}_G F \equiv F = G) \qquad \bowtie$$

(426.2) Definition (421) implies:

$$\mathit{ThinFormOf}(\boldsymbol{a}_G, G) \;\equiv\; A!\boldsymbol{a}_G \,\&\, G\!\downarrow \,\&\, \forall F(\boldsymbol{a}_G F \equiv F = G)$$

By Rule ≡S of Biconditional Simplification and the fact that $G\!\downarrow$ is an axiom, it follows that:

$$\mathit{ThinFormOf}(\boldsymbol{a}_G, G) \;\equiv\; A!\boldsymbol{a}_G \,\&\, \forall F(\boldsymbol{a}_G F \equiv F = G)$$

Hence, by (426.1), $\mathit{ThinFormOf}(\boldsymbol{a}_G, G)$. ⋈

(427) Assume $\mathit{ThinFormOf}(x, G)$. Then by definition of *ThinFormOf* (421), it follows *a fortiori* that $\forall F(xF \equiv F = G)$. Instantiating to G, we have $xG \equiv G = G$. But the right side is known, by the special case of Rule =I (118.2). Hence xG. ⋈

(429) Assume $\mathit{ThinFormOf}(x, G)$. By GEN, we show $Gy \equiv \mathit{ParticipatesIn}(y, x)$. (→) Assume Gy. Then by conjoining our two assumptions and applying ∃I, it follows that $\exists F(\mathit{ThinFormOf}(x, F) \,\&\, Fy)$. So by the definition of *ParticipatesIn* (428), it follows that $\mathit{ParticipatesIn}(y, x)$. (←) Assume $\mathit{ParticipatesIn}(y, x)$. By definition of *ParticipatesIn* (428), it follows that $\exists F(\mathit{ThinFormOf}(x,F) \& Fy)$. Assume P is an arbitrary such property, so that we know both (a) $\mathit{ThinFormOf}(x,P)$ and (b) Py. From (a), it follows that xP, by (427). But our global assumption is $\mathit{ThinFormOf}(x,G)$, from which it follows by definition of *ThinFormOf* (421) that $\forall F(xF \equiv F = G)$. Hence $xP \equiv P = G$. Since we've established xP, it follows that $P = G$. So from (b) it follows that Gy, by Rule =E. ⋈

(430) Since $\boldsymbol{a}_G$ exists, we can instantiate (429) to obtain:

$$\mathit{ThinFormOf}(\boldsymbol{a}_G, G) \rightarrow \forall y(Gy \equiv \mathit{ParticipatesIn}(y, \boldsymbol{a}_G))$$

But we know the antecedent by (426.2). Hence:

$$\forall y(Gy \equiv \mathit{ParticipatesIn}(y, \boldsymbol{a}_G))$$

By $\forall$E, $Gx \equiv \mathit{ParticipatesIn}(x, \boldsymbol{a}_G)$. ⋈

(431) Assume $Gx \& Gy \& x \neq y$. Since $\boldsymbol{a}_G$ exists, it follows by Rule =I (118.1) that $\boldsymbol{a}_G = \boldsymbol{a}_G$. Independently, it follows from the first conjunct of our assumption and (430) that $\mathit{Participates}(x, \boldsymbol{a}_G)$. By similar reasoning from the second conjunct of our assumption, we can derive $\mathit{Participates}(y, \boldsymbol{a}_G)$. Hence, we have established that:

$$\boldsymbol{a}_G = \boldsymbol{a}_G \,\&\, \mathit{Participates}(x, \boldsymbol{a}_G) \,\&\, \mathit{Participates}(y, \boldsymbol{a}_G)$$

Hence, $\exists z(z = \boldsymbol{a}_G \,\&\, \mathit{ParticipatesIn}(x,z) \,\&\, \mathit{ParticipatesIn}(y,z))$. ⋈

(432.1) – **(432.2)** (Exercises)

(433.1) By (426.2), (427), and the fact that $\boldsymbol{a}_G$ exists. ⋈

(433.2) By the second conjunct of (426.1), we know $\forall F(\boldsymbol{a}_G F \equiv F = G)$. Hence $\exists H \forall F(\boldsymbol{a}_G F \equiv F = H))$, i.e., $\exists! H \boldsymbol{a}_G H$, by (127.2). ⋈

(436) (Exercise)

(437.1) Assume $\mathit{ThinForm}(x)$. Then by definition (435.1), $\exists F(\mathit{ThinFormOf}(x,F))$. Assume P is an arbitrary such property, so that we know $\mathit{ThinFormOf}(x,P)$. Then by applying definition (421), it follows that $A!x$. ⋈

(437.2) Assume $\Box\forall y(A!y \rightarrow Fy)$ and $\mathit{ThinForm}(x)$. Our second assumption implies $A!x$, by (437.1). Independently, our first assumption and the T schema imply that $\forall y(A!y \rightarrow Fy)$, and so $A!x \rightarrow Fx$. Hence Fx. ⋈

(437.3) By instantiating $A!$ for G in the left conjunct of (426.1). ⋈

(437.4) Assume $\mathit{ThinForm}(x)$. Then by (437.1), it follows that $A!x$. But then by (430), it follows that $\mathit{Participates}(x, \boldsymbol{a}_{A!})$. ⋈

(437.5) Assume $\Box\forall y(A!y \rightarrow Fy)$ and $ThinForm(x)$. Then by (437.2), it follows from our two assumptions that Fx. By (430), it follows that $ParticipatesIn(x, \boldsymbol{a}_F)$. ⋈

(438.1) By (436), $ThinForm(\boldsymbol{a}_{A!})$. So by (437.4), $ParticipatesIn(\boldsymbol{a}_{A!}, \boldsymbol{a}_{A!})$. By &I and ∃I, we're done. ⋈

(438.2) By (436), $ThinForm(\boldsymbol{a}_{O!})$. So by (430), $\forall y(O!y \equiv ParticipatesIn(y, \boldsymbol{a}_{O!}))$. Instantiating to $\boldsymbol{a}_{O!}$, we have: $O!\boldsymbol{a}_{O!} \equiv ParticipatesIn(\boldsymbol{a}_{O!}, \boldsymbol{a}_{O!})$. But by (432.2), we know $\neg O!\boldsymbol{a}_{O!}$. Hence, $\neg ParticipatesIn(\boldsymbol{a}_{O!}, \boldsymbol{a}_{O!})$. By &I and ∃I, we're done. ⋈

(440.1) For reductio, assume:

$$[\lambda x\, ThinForm(x) \,\&\, \neg ParticipatesIn(x,x)]\downarrow$$

To simplify notation, let's abbreviate the λ-expression as K, so that our assumption is $K\downarrow$. Then by (422.1), $\exists x ThinFormOf(x, K)$. Let a be such an object, so that we know $ThinFormOf(a, K)$. Moreover, given the definition of K and the fact that $K\downarrow$, it follows by β-Conversion that:

(ϑ) $Ka \equiv ThinForm(a) \,\&\, \neg ParticipatesIn(a,a)$

Now we can conclude our reductio if we can establish either of two contradictory claims:

$$Ka \equiv \neg Ka$$

$$ParticipatesIn(a,a) \equiv \neg ParticipatesIn(a,a)$$

We'll establish the first and leave the second as an exercise. (→) Take Ka as a local assumption. Independently, we may appeal to (429) to infer from the fact that $ThinFormOf(a, K)$ that $\forall y(Ky \equiv ParticipatesIn(y, a))$. From this and our local assumption it follows that $ParticipatesIn(a, a)$. But given our local assumption, it follows from (ϑ) *a fortiori* that $\neg ParticipatesIn(a,a)$. Contradiction. Hence $\neg Ka$.

(←) Now take $\neg Ka$ as a local assumption. Then by (ϑ), $\neg(ThinForm(a) \,\&\, \neg ParticipatesIn(a,a))$, i.e.,

(ξ) $ThinForm(a) \rightarrow ParticipatesIn(a,a)$

But from $ThinFormOf(a, K)$ it follows that $\exists G(ThinFormOf(a, G))$ and so by definition (435), $ThinForm(a)$. From this and (ξ), we have $ParticipatesIn(a,a)$. So by definition (428), $\exists G(ThinFormOf(a, G) \,\&\, Ga)$. Let P be such a property, so that we know both $ThinFormOf(a, P)$ and Pa. But by (422.4), it follows from $ThinFormOf(a, P)$ and $ThinFormOf(a, K)$ that $P = K$. Hence, Ka. ⋈

(440.2) (Exercise)

(443.1) – **(443.2)** (Exercise)

(443.3) (→) Assume $G \Leftrightarrow F$. By definitions (442.1) and (442.2), we know *a fortiori*:

(a) $\Box\forall x(Gx \to Fx)$

(b) $\Box\forall x(Fx \to Gx)$

Since H doesn't have a free occurrence in our assumption, it suffices by GEN to show $G \Rightarrow H \equiv F \Rightarrow H$:

> (→) Assume $G \Rightarrow H$. By (442.1), we know *a fortiori* that $H\downarrow$ and $\Box\forall x(Gx \to Hx)$. So from (b) and this latter conclusion, it follows that $\Box\forall x(Fx \to Hx)$, by (168.5). Since $F\downarrow$ is axiomatic, it follows that $F \Rightarrow H$ by definition (442.1).
>
> (←) By analogous reasoning using (a).

(←) Assume $\forall H(G \Rightarrow H \ \equiv F \Rightarrow H)$. By definition of $\Leftrightarrow$ (442), we have to show both (a) $G \Rightarrow F$ and (b) $F \Rightarrow G$. To show (a), instantiate G into our assumption, to obtain $G \Rightarrow G \ \equiv F \Rightarrow G$. The left condition follows from the fact that $\Leftrightarrow$ is an equivalence condition and so reflexive (443.2.a). Hence, $F \Rightarrow G$. To show (b), reason analogously by instantiating our assumption to F. ⋈

(443.4) (Exercise)

(443.5) Before we begin, note two things. First that by Rule ≡S of Biconditional Simplification and definition (442.1), we know the following is a modally strict theorem:

(A) $F \Rightarrow H \equiv \Box\forall x(Fx \to Hx)$

Second, the following lemma holds:

(B) $\Box\forall x\neg Fx \ \to \ F \Rightarrow H$

Proof:

1.	$\neg Fx \to (Fx \to Hx)$	instance of (77.3)
2.	$\forall x(\neg Fx \to (Fx \to Hx))$	from 1, by GEN
3.	$\forall x\neg Fx \to \forall x(Fx \to Hx)$	from 2, by axiom (39.3)
4.	$\Box(\forall x\neg Fx \to \forall x(Fx \to Hx))$	from 3, by RN
5.	$\Box\forall x\neg Fx \to \Box\forall x(Fx \to Hx)$	from 4, by K (45.1)
6.	$\Box\forall x\neg Fx \ \to \ F \Rightarrow H$	from 5, by (A) and substitution

Now to prove our theorem, assume *Impossible*(G) & *Impossible*(F). Hence, we know:

(C) $\Box\forall x \neg Gx$

(D) $\Box\forall x \neg Fx$

Now by (443.3) and GEN, it suffices to show $G \Rightarrow H \equiv F \Rightarrow H$. ($\rightarrow$) From (B) and (D), it follows that $F \Rightarrow H$. Hence, $G \Rightarrow H \rightarrow F \Rightarrow H$, by axiom (38.1). ($\leftarrow$) By GEN, (B) implies $\forall F \forall H(\Box\forall x \neg Fx \rightarrow F \Rightarrow H)$ by GEN. Hence instantiating to G and H, we obtain $\Box\forall x \neg Gx \rightarrow G \Rightarrow H$. From this and (C), it follows that $G \Rightarrow H$. Hence $F \Rightarrow H \rightarrow G \Rightarrow H$, by axiom (38.1). ⋈

(445.1) – **(445.3)** (Exercises)

(447) (Exercise)

(448) By GEN and RN, it suffices to show $G \Rightarrow F \rightarrow \Box G \Rightarrow F$. So assume $G \Rightarrow F$. Independently, from definition (442.1) and the facts that $G\downarrow$ and $F\downarrow$, the following is a modally strict theorem, by Rule ≡S of Biconditional Simplification that:

(ϑ) $G \Rightarrow F \equiv \Box\forall x(Gx \rightarrow Fx)$

Hence. $\Box\forall x(Gx \rightarrow Fx)$. So by the 4 schema, $\Box\Box\forall x(Gx \rightarrow Fx)$. Hence from ($\vartheta$) and a Rule of Substitution (160.2), it follows that $\Box G \Rightarrow F$. ⋈

(449.1) By (448) and (260.1), we know that $G \Rightarrow F$ is a rigid condition on properties. So the following is an instance of (261.2):

$$y = \imath x(A!x \,\&\, \forall F(xF \equiv G \Rightarrow F)) \rightarrow (A!y \,\&\, \forall F(yF \equiv G \Rightarrow F))$$

By GEN, the fact that $\Phi_G\downarrow$, and $\forall$E, it follows that:

$$\Phi_G = \imath x(A!x \,\&\, \forall F(xF \equiv G \Rightarrow F)) \rightarrow (A!\Phi_G \,\&\, \forall F(\Phi_G F \equiv G \Rightarrow F))$$

So by theorem (447), $A!\Phi_G \,\&\, \forall F(\Phi_G F \equiv G \Rightarrow F)$. ⋈

(449.2) From (449.1) and the fact that $G\downarrow$, by definition (444). ⋈

(452.1) Assume $FormOf(x, G)$. By GEN, it suffices to show:

$$Gy \equiv ParticipatesIn_{\text{PTA}}(y, x)$$

($\rightarrow$) Assume Gy. Then by conjoining our two assumptions and applying $\exists$I, it follows that $\exists F(FormOf(x, F) \,\&\, Fy)$. So by the definition of $ParticipatesIn_{\text{PTA}}$ (451.1), it follows that $ParticipatesIn_{\text{PTA}}(y, x)$. ($\leftarrow$) Assume $ParticipatesIn_{\text{PTA}}(y, x)$. By definition of $ParticipatesIn_{\text{PTA}}$ (451.1), it follows that $\exists F(FormOf(x, F) \,\&\, Fy)$. Assume P is an arbitrary such property, so that we know both (a) $FormOf(x, P)$ and (b) Py. From (a), it follows *a fortiori* by definition (444) that $\forall F(xF \equiv P \Rightarrow F)$. But we also know $FormOf(x, G)$, from which it also follows by (444) that $\forall F(xF \equiv G \Rightarrow F)$. By the laws of quantified biconditional, (99.11) and (99.10), it thus follows that $\forall F(P \Rightarrow F \equiv G \Rightarrow F)$. But by a fact about necessary equivalence (443.3), it follows that $P \Leftrightarrow G$, i.e., by (443.1), that $\Box\forall x(Px \equiv Gx)$. By the T schema, $\forall x(Px \equiv Gx)$, and by $\forall$E, $Py \equiv Gy$. But Py is already known. Hence Gy. ⋈

(452.2) Assume $FormOf(x,G)$. By GEN, it suffices to show:

$$yG \rightarrow ParticipatesIn_{\text{PH}}(y,x)$$

So assume yG. Then by conjoining our two assumptions and applying $\exists$I, it follows that $\exists F(FormOf(x,F) \,\&\, yF)$. So by the definition of $ParticipatesIn_{\text{PH}}$ (451.2), it follows that $ParticipatesIn_{\text{PH}}(y,x)$. ⋈

(454.1) Instantiate Φ_G into (452.1) to obtain:

$$FormOf(\Phi_G,G) \rightarrow \forall y(Gy \equiv ParticipatesIn_{\text{PTA}}(y,\Phi_G))$$

Then by (449.2), it follows that $\forall y(Gy \equiv ParticipatesIn_{\text{PTA}}(y,\Phi_G))$. Instantiate this result to x and we obtain: $Gx \equiv ParticipatesIn_{\text{PTA}}(x,\Phi_G)$. ⋈

(454.2) Instantiate Φ_G into (452.2) to obtain:

$$FormOf(\Phi_G,G) \rightarrow \forall y(yG \rightarrow ParticipatesIn_{\text{PH}}(y,\Phi_G))$$

Then by (449.2), it follows that, $\forall y(yG \rightarrow ParticipatesIn_{\text{PH}}(y,\Phi_G))$. Instantiate this result to x and we obtain: $xG \rightarrow ParticipatesIn_{\text{PH}}(x,\Phi_G)$. ⋈

(455) Assume $ParticipatesIn_{\text{PTA}}(y,x)$. Then $\exists F(FormOf(x,F) \,\&\, Fy)$, by (451.1). Suppose P is an arbitrary such property, so that we know both $FormOf(x,P)$ and Py. By GEN, it suffices to show $xF \rightarrow Fy$. So assume xF. Then $P \Rightarrow F$, by the established fact that $FormOf(x,P)$ and definition (444). So $\Box\forall y(Py \rightarrow Fy)$, by (442.1). Since Py is already known, it follows by now familiar reasoning that Fy. ⋈

(456.1) Assume $Gx \,\&\, Gy \,\&\, x \neq y$. Since Φ_G exists, it follows by Rule =I (118.1) that $\Phi_G = \Phi_G$. Independently, it follows from the first conjunct of our assumption and (454.1) that $ParticipatesIn_{\text{PTA}}(x,\Phi_G)$. By similar reasoning from the second conjunct of our assumption, we can derive $ParticipatesIn_{\text{PTA}}(y,\Phi_G)$. Hence, we have established that:

$$\Phi_G = \Phi_G \,\&\, ParticipatesIn_{\text{PTA}}(x,\Phi_G) \,\&\, ParticipatesIn_{\text{PTA}}(y,\Phi_G)$$

By $\exists$I, we're done. ⋈

(456.2) (Exercise)

(457.1) By reasoning analogous to the proof of (432.1).

(457.2) By reasoning analogous to the proof of (432.2).

(457.3) (Exercise)

(458) By instantiating G into the second conjunct of (449.1), it follows that $\Phi_G G \equiv G \Rightarrow G$. But the right-hand side, by definition (442.1), is just $\Box\forall x(Gx \rightarrow Gx)$, which is easily derivable by applying GEN and then RN to the tautology $Gx \rightarrow Gx$. So $\Phi_G G$. ⋈

(460.1) – **(460.3)** (Exercises)

(461.1) Instantiate the first conjunct of (449.1) to $A!$. ⋈

(461.2) By (461.1), we know $A!\Phi_{A!}$. It follows that $\mathit{ParticipatesIn}_{\text{PTA}}(\Phi_{A!}, \Phi_{A!})$, by (454.1). ⋈

(461.3) By (460.1), we know $\mathit{Form}(\Phi_{A!})$. Conjoin this with (461.2) and existentially generalize. ⋈

(461.4) Assume $A! \Rightarrow H$. Since G isn't free in our assumption, it suffices by GEN to show $H\Phi_G$. By definition (442), it follows from our assumption that $\Box\forall x(A!x \rightarrow Hx)$, and by the T schema, that $\forall x(A!x \rightarrow Hx)$. By the first conjunct of (449.1), we know $A!\Phi_G$. Hence $H\Phi_G$. ⋈

(461.5) (Exercise)

(461.6) Assume $A! \Rightarrow \overline{H}$. By GEN, it suffices to show $\neg H\Phi_G$. By (461.4), it follows from our assumption that $\forall G(\overline{H}\Phi_G)$, and hence $\overline{H}\Phi_G$. By theorem (199.1), it follows that $\neg H\Phi_G$. ⋈

(461.7) (Exercise)

(462.1) Assume $A! \Rightarrow H$. Then by (461.4), it follows that $\forall G(H\Phi_G)$. But since theorem (454.1) is modally strict, there is a modally strict proof of its instance $H\Phi_G \equiv \mathit{ParticipatesIn}_{\text{PTA}}(\Phi_G, \Phi_H)$. Hence, by the Rule of Substitution (160.2), $\forall G(\mathit{ParticipatesIn}_{\text{PTA}}(\Phi_G, \Phi_H))$. ⋈

(462.2) Assume $G \Rightarrow H$. Then by the second conjunct of (449.1), it follows that $\Phi_G H$. So by an appropriate instance of (454.1), obtained by substituting Φ_G for x and H for G, it follows that $\mathit{ParticipatesIn}_{\text{PH}}(\Phi_G, \Phi_H)$. ⋈

(467.2) – **(467.3)** (Exercises)

(468) Assume $\mathit{TruthValue}(x)$. Then by (290), $\exists p(\mathit{TruthValueOf}(x,p))$. Let p_1 be such a proposition, so that we know $\mathit{TruthValueOf}(x,p_1)$. Hence by (286), it follows that:

$$(\vartheta)\quad A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p_1) \,\&\, F = [\lambda y\, q]))$$

Now since the first conjunct of (ϑ) is that $A!x$, it remains to show, by definition (467), that $\forall F(xF \rightarrow \mathit{Propositional}(F))$. By GEN, it suffices to show $xF \rightarrow \mathit{Propositional}(F)$. So assume xF. Then by the second conjunct of (ϑ) it follows that:

$$\exists q((q \equiv p_1) \,\&\, F = [\lambda y\, q])$$

It follows *a fortiori* that $\exists q(F = [\lambda y\, q])$, for which $\exists p(F = [\lambda y\, p])$ is an alphabetic variant. So *Propositional*(F), by definition (275). ⋈

(469.1) ($\rightarrow$) Suppose *Situation*(x). Then, by the definition of *Situation* (467), we know:

$$(\vartheta)\ \ A!x \,\&\, \forall F(xF \rightarrow \exists p(F = [\lambda y\, p]))$$

To show $\Box(\vartheta)$, it suffices to show that both conjuncts of (ϑ) are necessary, by right-to-left direction of theorem (158.3), i.e., by the fact that $(\Box\varphi \,\&\, \Box\psi) \rightarrow \Box(\varphi \,\&\, \psi)$. But the first conjunct of (ϑ), i.e., $A!x$, implies $\Box A!x$, by (180.2). And the second conjunct of (ϑ) also implies its own necessity, by (281.2). ($\leftarrow$) Exercise. ⋈

(469.2) – **(469.3)** (Exercises)

(469.4) ($\rightarrow$) Assume $\mathcal{A}$*Situation*(x). Independently, since the left-to-right direction of (469.1) is *Situation*$(x) \rightarrow \Box$*Situation*(x), it follows by RA that:

$$\mathcal{A}(Situation(x) \rightarrow \Box Situation(x))$$

So by theorem (131):

$$\mathcal{A}Situation(x) \rightarrow \mathcal{A}\Box Situation(x)$$

Hence, $\mathcal{A}\Box$*Situation*(x), given our assumption. So $\Box$*Situation*(x), by (46.2). And thus *Situation*(x), by the T schema. ($\leftarrow$) Assume *Situation*(x). Then by (469.1), $\Box$*Situation*(x). So by (132), $\mathcal{A}$*Situation*(x). ⋈

(469.5) As noted in the text, this can be proved using the Rule of Modal Strictness (173) from the fact that *Situation*$(\bot)$ (which we mentioned was derivable at the end of (468)) and the fact that *Situation*(x) is modally collapsed (which is derivable from (469.1)). But what follows is a proof without the Rule of Modal Strictness.

By definition (467), we have to show:

$$A!{\circ}p \,\&\, \forall F({\circ}pF \rightarrow Propositional(F))$$

By theorem (296.1), we know:

$${\circ}p = \imath x(A!x \,\&\, \forall F(xF \equiv \exists q((q \equiv p) \,\&\, F = [\lambda y\, q])))$$

By (255), it follows that $A! \circ p$. So it remains to show ${\circ}pF \rightarrow Propositional(F)$, by GEN. Assume ${\circ}pF$. By definition (275), we have to show $\exists p(F = [\lambda y\, p])$. Independently, in the solution to the Exercise in (299) – see footnote 197 – we established that:

$$\circ pF \equiv \mathcal{A}\exists q((q \equiv p) \,\&\, F = [\lambda y\, q])$$

Hence $\mathcal{A}\exists q((q \equiv p) \,\&\, F = [\lambda y\, q])$. *A fortiori*, $\mathcal{A}\exists q(F = [\lambda y\, q])$ (exercise). So by by (139.10), $\exists q\mathcal{A}(F = [\lambda y\, q])$. Suppose q_1 is such a proposition, so that we know $\mathcal{A}(F = [\lambda y\, q_1])$. But by the right-to-left direction of (175.1), it follows that $F = [\lambda y\, q_1]$. Hence $\exists p(F = [\lambda y\, p])$. ⋈

(469.6) (Exercises)

(469.7) As noted in the text, this can be proved using the Rule of Modal Strictness (173) from the fact that $Situation(\bot)$, which we mentioned was derivable at the end of (468), and the fact that $Situation(x)$ is modally collapsed, which is derivable from (469.1).

But here is a proof without the Rule of Modal Strictness: By definition (467), we have to show:

$$A!\bot \,\&\, \forall F(\bot F \to Propositional(F))$$

By definition of $\bot$ (302.2) and (255), it follows that $A!\bot$. So, by GEN, it remains to show $\bot F \to Propositional(F)$. Assume $\bot F$. By definition (275), we have to show $\exists p(F = [\lambda y\, p])$. By (258.1) and the definition of $\bot$, we know $\bot F \equiv \mathcal{A}\exists p(\neg p \,\&\, F = [\lambda y\, p])$. Hence $\mathcal{A}\exists p(\neg p \,\&\, F = [\lambda y\, p])$. *A fortiori*, $\mathcal{A}\exists p(F = [\lambda y\, p])$ (exercise). So $\exists p\mathcal{A}(F = [\lambda y\, p])$, by (139.10). Suppose q_1 is such a proposition, so that we know $\mathcal{A}(F = [\lambda y\, q_1])$. But by the right-to-left direction of (175.1), it follows that $F = [\lambda y\, q_1]$. Hence $\exists p(F = [\lambda y\, p])$. ⋈

(471) Assume $Situation(x)$. ($\to$) Assume $x \models p$. Then it follows *a fortiori* from definition (470) that $x\Sigma p$. From this it follows from definition (295) that $x[\lambda y p]$. ($\leftarrow$) Assume $x[\lambda y\, p]$. So by definition (295), it follows that:

$$x\Sigma p \equiv x[\lambda y\, p]$$

So from this and our local assumption it follows that $x\Sigma p$. From our global assumption that $Situation(x)$ and this last result, it follows, by definition (470), that $x \models p$. ⋈

(473.1) If we eliminate the restricted variable, we have to prove: $Situation(x) \to (x \models p \equiv \Box x \models p)$. So assume $Situation(x)$.

($\to$) Assume $x \models p$. Then by (471), it follows that $x[\lambda y\, p]$. Since this is an encoding formula, axiom (51) applies and yields:

$$(\vartheta) \quad \Box x[\lambda y\, p]$$

But note also that theorem (471) is modally strict, and so by Rule RM, it implies:

$$\Box Situation(x) \to \Box((x \models p) \equiv x[\lambda y\, p])$$

But our global assumption is that $Situation(x)$, which implies $\Box Situation(x)$, by (469.1). Hence:

(ξ) $\Box((x \models p) \equiv x[\lambda y\, p])$

But from (ξ) and (ϑ), it follows that $\Box x \models p$, by a version of (158.6).

($\leftarrow$) By the T schema. ⋈

[Note: In light of Remark (472), and our previous discussions of reasoning with restricted variables in (340) and (341.2), we may use our the rigid restricted variable s to simplify the reasoning in the left-to-right direction, as follows. ($\rightarrow$) Assume $s \models p$. Then by (471), $s[\lambda y\, p]$. So by axiom (51), $\Box s[\lambda y\, p]$. But since the antecedent of (471) is, by (469.1), necessarily true if true, we can independently derive, using restricted variables, $\Box((s \models p) \equiv s[\lambda y\, p])$ (exercise). Hence $\Box s \models p$.]

(473.2) – **(473.4)** (Exercises)

(473.5) Theorem (473.2) is that $\Diamond s \models p \equiv s \models p$. So by a classical tautology (88.4.b), $\neg\Diamond s \models p \equiv \neg s \models p$. Independently, as an instance of (162.1), we know $\Box\neg s \models p \equiv \neg\Diamond s \models p$. So by transitivity, $\Box\neg s \models p \equiv \neg s \models p$, and by commutativity, $\neg s \models p \equiv \Box\neg s \models p$. ⋈

(474) ($\rightarrow$) Exercise. ($\leftarrow$) Assume $\forall p(s \models p \equiv s' \models p)$. Since both s and s' are situations, it follows that they are abstract objects, by definition (467). By (245.2), it suffices to show that s and s' encode the same properties:

> ($\rightarrow$) Assume sF. Then since s is a situation, it follows that $\exists p(F = [\lambda y\, p])$, by definitions (467) and (275). Suppose p_1 is an arbitrary such proposition, so that we know $F = [\lambda y\, p_1]$. So $s[\lambda y\, p_1]$ and by (471), $s \models p_1$. But our initial hypothesis is that the same propositions are true in s and s'. So $s' \models p_1$. Hence by (471), $s'[\lambda y\, p_1]$. So $s'F$.
>
> ($\leftarrow$) By analogous reasoning. ⋈

(476.1) (Exercise)

(476.2) Assume $s \trianglelefteq s'$ and $s \neq s'$. For reductio, assume $s' \trianglelefteq s$. From the assumption that $s \trianglelefteq s'$, it follows that $\forall p(s \models p \rightarrow s' \models p)$, by definition (475). Similarly, from the reductio assumption, it follows that $\forall p(s' \models p \rightarrow s \models p)$. Hence, $\forall p(s \models p \equiv s' \models p)$. So by theorem (474), $s = s'$. Contradiction. ⋈

(476.3) Assume (a) $s \trianglelefteq s'$ and (b) $s' \trianglelefteq s''$. To show $s \trianglelefteq s''$, assume $s \models p$. From this and (a), it follows by definition of $\trianglelefteq$ (475) that $s' \models p$. From this and (b), it follows by (475) that $s'' \models p$. ⋈

(477.1) ($\rightarrow$) Exercise. ($\leftarrow$) By exportation (88.7.a), the anti-symmetry of $\trianglelefteq$ (476.2) becomes $s \trianglelefteq s' \rightarrow (s \neq s' \rightarrow \neg(s' \trianglelefteq s))$. By contraposition of the consequent, we get $s \trianglelefteq s' \rightarrow (s' \trianglelefteq s \rightarrow s = s')$. So by importation (88.7.b), $(s \trianglelefteq s' \,\&\, s' \trianglelefteq s) \rightarrow s = s'$. ⋈

(477.2) ($\rightarrow$) Exercise. ($\leftarrow$) We take $\forall s''(s'' \trianglelefteq s \equiv s'' \trianglelefteq s')$ as a global assumption. Given theorem (474), it suffices to show $\forall p(s \models p \equiv s' \models p)$. By GEN, it suffices to show $s \models p \equiv s' \models p$:

> ($\rightarrow$) Assume $s \models p$. By instantiating our global assumption to s, we know $s \trianglelefteq s \equiv s \trianglelefteq s'$. But by (476.1), we know $s \trianglelefteq s$. Hence $s \trianglelefteq s'$. So by definition of $\trianglelefteq$ (475), it follows that $s' \models p$.
>
> ($\leftarrow$) Assume $s' \models p$. Then $s \models p$ follows by analogous reasoning once we instantiate our global assumption to s' and let our instance of (476.1) be $s' \trianglelefteq s'$. ⋈

(479) (Exercise)

(481) Suppose φ is a condition on propositional properties. Then by (480), we know there is a modally strict proof of:

(ϑ) $\forall F(\varphi \rightarrow \mathit{Propositional}(F))$

Now if we let θ be $\mathit{Situation}(x)$, ψ be $\forall F(xF \equiv \varphi)$, and χ be $A!x$, then the theorem we have to prove has the form:

$$(\theta \,\&\, \psi) \equiv (\chi \,\&\, \psi)$$

But to prove this, it suffices to show $\psi \rightarrow (\theta \equiv \chi)$, in light of the right-to-left direction of the tautology (88.8.f):

$$((\theta \,\&\, \psi) \equiv (\chi \,\&\, \psi)) \equiv (\psi \rightarrow (\theta \equiv \chi))$$

So we want to show $\forall F(xF \equiv \varphi) \rightarrow (\mathit{Situation}(x) \equiv A!x)$. Assume:

(ζ) $\forall F(xF \equiv \varphi)$

($\rightarrow$) Assume $\mathit{Situation}(x)$. But $A!x$ follows immediately by definition (467) of *Situation*.

($\leftarrow$) Assume $A!x$. In virtue of the definition (467) of *Situation*, it remains only to show $\forall F(xF \rightarrow \mathit{Propositional}(F))$. By GEN, we need only show: $xF \rightarrow \mathit{Propositional}(F)$. So assume xF. Then by (ζ), it follows that φ. But if we instantiate (ϑ) to F, it follows that $\varphi \rightarrow \mathit{Propositional}(F)$. Hence, $\mathit{Propositional}(F)$. ⋈

(482.1) Suppose φ is a condition on propositional properties in which x doesn't occur free. Since x doesn't occur free in φ, it follows by the Comprehension Principle for Abstract Objects (53) that:

$$\exists x(A!x \,\&\, \forall F(xF \equiv \varphi))$$

Since φ is a condition on propositional properties, it follows from the modally strict theorem (481) and the Rule of Substitution (160.2) that:

$$\exists x(Situation(x) \,\&\, \forall F(xF \equiv \varphi))$$ ⋈

(482.2) (Exercise)

(483.1) Let φ be a condition on propositional properties in which x doesn't occur free. Then by (482.2), it follows that:

$$\exists!x(Situation(x) \,\&\, \forall F(xF \equiv \varphi))$$

If apply the Rule of Actualization and then use the right-to-left direction of (176.2), it follows that:

$$\imath x(Situation(x) \,\&\, \forall F(xF \equiv \varphi))\downarrow$$ ⋈

(483.2) Let φ be a condition on propositional properties in which x doesn't occur free. Now by applying GEN and RN to (481), we know:

(ϑ) $\Box\forall x((Situation(x) \,\&\, \forall F(xF \equiv \varphi)) \equiv (A!x \,\&\, \forall F(xF \equiv \varphi)))$

Consequently, our theorem follows from (483.1) and (ϑ) by theorem (149.3). ⋈

(484) Suppose φ is a rigid condition on propositional properties in which x isn't free. Given our conventions for interpreting bound restricted variables, we have to show:

$$y = \imath x(Situation(x) \,\&\, \forall F(xF \equiv \varphi)) \to \forall F(yF \equiv \varphi))$$

So assume:

(ϑ) $y = \imath x(Situation(x) \,\&\, \forall F(xF \equiv \varphi))$

Since φ is, by hypothesis, a condition on propositional properties in which x doesn't occur free, we may appeal to the identity (483.2) and infer from (ϑ) that:

(ξ) $y = \imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$

But since φ is also, by hypothesis, a rigid condition on properties in which x doesn't occur free, it follows from (ξ) by (261.2) that:

$$A!y \,\&\, \forall F(yF \equiv \varphi)$$

A fortiori, $\forall F(yF \equiv \varphi)$. ⋈

(486.1) If we eliminate the restricted variable, then we have to show:

$$\exists x(Situation(x) \,\&\, \forall p(x \models p \equiv \varphi)), \text{ provided } x \text{ doesn't occur free in } \varphi$$

Now let φ be any formula in which x doesn't occur free and pick a property variable that doesn't occur free in φ. Without loss of generality, suppose F doesn't occur free in φ. Then, the formula $\varphi \,\&\, F = [\lambda y\, p]$ is nevertheless a condition on propositional properties (480) – it is provable, by modally strict means, that any property that satisfies this condition is a propositional property. Hence, by (482.1), we know:[464]

$$\exists x(Situation(x) \,\&\, \forall F(xF \equiv \varphi \,\&\, F = [\lambda y\, p]))$$

Let a be such a situation, so that we know:

$$(\vartheta)\quad Situation(a) \,\&\, \forall F(aF \equiv \varphi \,\&\, F = [\lambda y\, p])$$

It then remains only to show $\forall p(a \models p \equiv \varphi)$. Since we've not made any special assumptions about p other than citing theorems in which p occurs free, we need only show $a \models p \equiv \varphi$, by GEN. ($\rightarrow$) Assume $a \models p$. Then since a is a situation, it follows by (471) that $a[\lambda y\, p]$. Hence, by the second conjunct of (ϑ) and the fact that F doesn't occur free in φ, it follows that $\varphi \,\&\, [\lambda y\, p] = [\lambda y\, p]$.[465] *A fortiori*, φ. ($\leftarrow$) Assume φ. Then by =I and &I, $\varphi \,\&\, [\lambda y\, p] = [\lambda y\, p]$. Hence, by the second conjunct of (ϑ), $a[\lambda y\, p]$. So by (471) and the fact that a is a situation, $a \models p$. ⋈

(486.2) – **(486.4)** (Exercises)

(488.1) By the simplified comprehension principle for situations (486.1), we know:

$$\exists s \forall p(s \models p \equiv p \neq p)$$

If we eliminate the restricted variable, this becomes:

$$\exists x(Situation(x) \,\&\, \forall p(x \models p \equiv p \neq p))$$

[464] If φ is a formula in which x doesn't occur free but F does occur free, then pick some property variable, say G, that doesn't occur free in φ. Then the following is an instance of comprehension for situations:

$$\exists x(Situation(x) \,\&\, \forall G(xG \equiv \varphi \,\&\, G = [\lambda y\, p]))$$

Now use this alphabetic variant in the proof that follows.

[465] To be maximally explicit, when we instantiate $a[\lambda y\, p]$ into the universal claim $\forall F(aF \equiv \varphi \,\&\, F = [\lambda y\, p])$, we may infer the result of substituting $[\lambda y\, p]$ for all the free occurrences of F in the formula $aF \equiv \varphi \,\&\, F = [\lambda y\, p]$), which is:

$$a[\lambda y\, p] \equiv \varphi^{[\lambda y\, p]}_{F} \,\&\, [\lambda y\, p] = [\lambda y\, p]$$

Since we've established $a[\lambda y\, p]$, it follows that:

$$\varphi^{[\lambda y\, p]}_{F} \,\&\, [\lambda y\, p] = [\lambda y\, p]$$

But since F doesn't occur free in φ, the first conjunct is just φ. So we reach the conclusion just mentioned in the text, namely, $\varphi \,\&\, [\lambda y\, p] = [\lambda y\, p]$.

Suppose x_0 is an arbitrary such object, so that we know:

(ξ) $Situtation(x_0) \,\&\, \forall p(x_0 \models p \equiv p \neq p)$

By definition of $NullSituation(x)$ (487.1), it remains to show:

(a) $\neg\exists p(x_0 \models p)$

(b) $\forall y(Situation(y) \,\&\, \neg\exists p(y \models p)) \rightarrow y = x_0)$

(a) Suppose for reductio that some proposition, say q_1, is such that $x_0 \models q_1$. Then by (ξ), it follows that $q_1 \neq q_1$, which contradicts the fact that $q_1 = q_1$.

(b) Assume $Situation(y)$ and $\neg\exists p(y \models p)$. By (a), we know $\neg\exists p(x_0 \models p)$. So it follows that $\forall p(y \models p \equiv x_0 \models p)$, by (103.9). Since both y and x_0 are situations, it follows by (474) that $y = x_0$. ⋈

(488.2) (Exercise)

(488.3) By the Rule of Actualization, (488.1) implies $\mathcal{A}\exists!xNullSituation(x)$. So by the right-to-left direction of (176.2), it follows that $\imath xNullSituation(x)\downarrow$. ⋈

(488.4) (Exercise)

(490.1) Assume $NullSituation(x)$. Then by definition (487.1), we know both:

(ϑ) $Situation(x)$

(ξ) $\neg\exists p(x \models p)$

Now to show $\Box NullSituation(x)$, we have to show:

$$\Box(Situation(x) \,\&\, \neg\exists p(x \models p))$$

By &I and (158.3), it suffices to show:

(a) $\Box Situation(x)$

(b) $\Box\neg\exists p(x \models p)$

(a) By the left-to-right direction of (469.1), (ϑ) implies $\Box Situation(x)$.

(b) Suppose, for reductio, $\neg\Box\neg\exists p(x \models p)$. Then by definition (18.5), $\Diamond\exists p(x \models p)$. So by BF$\Diamond$ (167.3), $\exists p \Diamond x \models p$. Suppose p_1 is an arbitrary such proposition, so that we know $\Diamond x \models p_1$. But by (ϑ), x is a situation. So by the left-to-right direction of (473.2), it follows that $x \models p_1$. But then, $\exists p(x \models p)$, by $\exists$I, which contradicts (ξ). ⋈

(490.2) (Exercise)

(490.3) Let ψ be the formula $NullSituation(x)$. As an instance of (153.2), we know:

$$\forall x(\psi \to \Box\psi) \to (\exists!x\psi \to \forall y(y = \imath x\psi \to \psi^y_x))$$

By applying GEN to (490.1), we know $\forall x(\psi \to \Box\psi)$. Hence:

$$\exists!x\psi \to \forall y(y = \imath x\psi \to \psi^y_x)$$

But (488.1) is $\exists!x\psi$. Hence:

$$\forall y(y = \imath x\psi \to \psi^y_x)$$

Since $\boldsymbol{s}_\varnothing\!\downarrow$, it follows that:

$$\boldsymbol{s}_\varnothing = \imath x\psi \to \psi^{\boldsymbol{s}_\varnothing}_x$$

Hence by definition (489.1), $\psi^{\boldsymbol{s}_\varnothing}_x$, i.e., $NullSituation(\boldsymbol{s}_\varnothing)$. ⋈

(490.4) (Exercise)

(491.1) (Exercise)

(491.2) By GEN and the Rule of Actualization, (491.1) implies:

$$\mathcal{A}\forall x(NullSituation(x) \equiv Null(x))$$

So by (149.1), it follows that:

$$\forall x(x = \imath xNullSituation(x) \;\equiv\; x = \imath xNull(x))$$

Since $\boldsymbol{s}_\varnothing\!\downarrow$, it follows that:

$$\boldsymbol{s}_\varnothing = \imath xNullSituation(x) \;\equiv\; \boldsymbol{s}_\varnothing = \imath xNull(x)$$

By definition (489.1), it follows that:

$$\boldsymbol{s}_\varnothing = \imath xNull(x)$$

Hence, by definition (265.1), $\boldsymbol{s}_\varnothing = \boldsymbol{a}_\varnothing$. ⋈

(491.3) Before we begin, note that by theorem (266.4), we know $Universal(\boldsymbol{a_V})$, and so by definition (263.2), that:

(ϑ) $A!\boldsymbol{a_V} \;\&\; \forall F\, \boldsymbol{a_V}F$

Furthermore, by theorem (490.4), we know $Trivial\ Situation(\boldsymbol{s_V})$, and so by definition (487.2), that:

(ζ) $Situation(\boldsymbol{s_V}) \;\&\; \forall p(\boldsymbol{s_V} \models p)$

Since the first conjunct of (ζ) implies $A!\boldsymbol{s_V}$, we've established that $\boldsymbol{a_V}$ and $\boldsymbol{s_V}$ are abstract. Hence, to establish our theorem, it suffices to show:

$$\exists F(\boldsymbol{a_V}F \;\&\; \neg\boldsymbol{s_V}F)$$

by (245.3). So we have to find a witness to this existential claim. If we take the witness to be $A!$, we have to show:

$$\boldsymbol{a_V} A! \,\&\, \neg \boldsymbol{s_V} A!$$

Now the first conjunct follows immediately by instantiating the second conjunct of (ϑ) to $A!$. So it remains to show the second conjunct. For reductio, suppose $\boldsymbol{s_V} A!$. Now since $\boldsymbol{s_V}$ is a situation, every property it encodes is a propositional property. So $\exists p(A! = [\lambda y\, p])$. Let p_1 be an arbitary such proposition, so that we know $A! = [\lambda y\, p_1]$. Independently, it is clear that $\exists x A!x \,\&\, \exists x \neg A!x$:

> $\exists x A!x$ follows from the Comprehension Principle for Abstract Objects. To show $\exists x \neg A!x$, it suffices, by the modally strict fact (222.2) and the Rule of Substitution (160.2), to show $\exists x O!x$. But this follows from (227.1), by the T schema (45.2).

So from $\exists x A!x \,\&\, \exists x \neg A!x$ and our identity $A! = [\lambda y\, p_1]$, it follows that

$$\exists x([\lambda y\, p_1]x) \,\&\, \exists x(\neg[\lambda y\, p_1]x)$$

But this yields a contradiction, for suppose b and c are arbitrary such an objects, so that we know both $[\lambda y\, p_1]b$ and $\neg[\lambda y\, p_1]c$. Now, clearly, $[\lambda y\, p_1]\downarrow$. So by β-Conversion, it follows from $[\lambda y\, p_1]b$ that p_1 and from $\neg[\lambda y\, p_1]c$ that $\neg p_1$. Contradiction. ⋈

(493) Theorem (217.1) tells us that there are contingently true propositions. So, by definition (213), we know $\exists p(p \,\&\, \Diamond\neg p)$. Suppose q_1 is such a proposition, so that we know:

(ϑ) $q_1 \,\&\, \Diamond\neg q_1$

Now consider the following instance of the simplified comprehension conditions for situations (486.1):

$$\exists s \forall p(s \models p \equiv p = q_1)$$

Suppose s_1 is such a situation, so that we know:

(ξ) $\forall p(s_1 \models p \equiv p = q_1)$

We now show that s_1 is a witness that proves our theorem by establishing:

(A) $Actual(s_1)$

(B) $\Diamond\neg Actual(s_1)$

(A) To show $Actual(s_1)$, we have to show $\forall p(s_1 \models p \rightarrow p)$. So by GEN, assume $s_1 \models p$. Then by (ξ), $p = q_1$. But by (ϑ), we know q_1. Hence p.

(B) Assume, for reductio, that $\neg\Diamond\neg Actual(s_1)$. Then, by (158.12), $\Box Actual(s_1)$, and so $\Box\forall p(s_1 \models p \rightarrow p)$, by definition of *Actual* (492) and a rule of substitution (160.3). By CBF (167.2), it follows that $\forall p \Box(s_1 \models p \rightarrow p)$. So, in particular:

(ζ) $\Box(s_1 \models q_1 \rightarrow q_1)$

But by (ξ), we know $s_1 \models q_1 \equiv q_1 = q_1$. Since the right side is true by the laws of identity, it follows that $s_1 \models q_1$. So by the rigidity of of truth in a situation (473.1), $\Box s_1 \models q_1$. But by the K axiom, this last conclusion and (ζ) together imply $\Box q_1$, i.e., $\neg\Diamond\neg q_1$, which contradicts the second conjunct of (ϑ). ⋈

(495.1) There are a number of ways to prove this theorem. Consider, for example, the null situation $\boldsymbol{s}_\varnothing$. Since no propositions are true in $\boldsymbol{s}_\varnothing$, by (490.3) and (487.1), it follows, by failure of the antecedent, that every proposition true in $\boldsymbol{s}_\varnothing$ is true.

But we can also easily construct a non-null actual situation. Consider the following instance of simplified comprehension for situations (486.1):

$$\exists s \forall p(s \models p \equiv p)$$

Let s_2 be an arbitrary such situation, so that we know $\forall p(s_2 \models p \equiv p)$. *A fortiori*, $(s_2 \models p) \rightarrow p$. Hence *Actual*($s_2$), by definition (492). ⋈

(495.2) As a witness to prove this claim, consider the trivial situation $\boldsymbol{s_V}$. By theorem (490.4) and definition (487.2), every proposition is true in $\boldsymbol{s_V}$. Hence, $\boldsymbol{s_V} \models (p \,\&\, \neg p)$, where p is any proposition you please. For reductio, suppose *Actual*($\boldsymbol{s_V}$). Then by definition of *Actual*, $p \,\&\, \neg p$. Contradiction. Hence $\neg$*Actual*($\boldsymbol{s_V}$). ⋈

(495.3) Consider the (necessarily) false proposition $\exists q(q \,\&\, \neg q)$, which we know exists. It suffices to show that this is a witness to our theorem. And by GEN, it suffices to show *Actual*$(s) \rightarrow \neg s \models \exists q(q \,\&\, \neg q)$. So assume *Actual*$(s)$. Then, by definition (492):

(ϑ) $\forall p(s \models p \rightarrow p)$

Now suppose, for reductio, that $s \models \exists q(q \,\&\, \neg q)$. Then by ($\vartheta$), it follows that $\exists q(q \,\&\, \neg q)$. Contradiction. ⋈

(496) By simplified comprehension for situations (486.1):

$$\exists s \forall p(s \models p \equiv s' \models p \vee s'' \models p)$$

Let s_3 be an arbitrary such situation, so that we know:

(ϑ) $\forall p(s_3 \models p \equiv s' \models p \vee s'' \models p)$

By &I and $\exists$I, we have to show (a) $s' \trianglelefteq s_3$, (b) $s'' \trianglelefteq s_3$, and (c) $\forall s'''(s' \trianglelefteq s''' \,\&\, s'' \trianglelefteq s''' \rightarrow s_3 \trianglelefteq s''')$. (a) By definition of $\trianglelefteq$ (475), we have to show $\forall p(s' \models p \rightarrow s_3 \models p)$. But this follows *a fortiori* from (ϑ). (b) By analogous reasoning. (c) By GEN, we want to show: $s' \trianglelefteq s''' \,\&\, s'' \trianglelefteq s''' \rightarrow s_3 \trianglelefteq s'''$. So let $s' \trianglelefteq s'''$ and $s'' \trianglelefteq s'''$ be our global assumptions. Since we want to show $\forall p(s_3 \models p \rightarrow s''' \models p)$ and p doesn't occur free in any assumption, it suffices to show $(s_3 \models p) \rightarrow (s''' \models p)$. So assume $s_3 \models p$. Then it follows from (ϑ) that $s' \models p \vee s'' \models p$. But the first disjunct implies $s''' \models p$ by our first global assumption, and the second disjunct implies $s''' \models p$ by our second global assumption. So $s''' \models p$, reasoning by cases. ⋈

(497.1) Assume *Actual*(s) and $s \models p$. By definition of *Actual* (492), it follows that p is true. Note independently that since $[\lambda y\ p]\downarrow$, we know by β-Conversion that $[\lambda y\ p]x \equiv p$. Since this last fact is a theorem, it follows by GEN that $\forall x([\lambda y\ p]x \equiv p)$. In particular, $[\lambda y\ p]s \equiv p$. But since we've established p, it follows that $[\lambda y\ p]s$. ⋈

(497.2) (Exercise) [Hint: Use reasoning analogous to the proof of (496).]

(500) Assume *Actual*(s). Then by definition (492):

(ϑ) $\forall p(s \models p \rightarrow p)$

Now suppose, for reductio, that $\neg$*Consistent*(s). So by Rule $\neg\neg$E (78.2) and the definition of *Consistent* (498), there is a proposition q, say q_1, such that both $s \models q_1$ and $s \models \neg q_1$. Then, it follows from the first and (ϑ) that q_1 and it follows from the second and (ϑ) that $\neg q_1$. Contradiction. ⋈

(501.1) $(\rightarrow)$ Assume $\neg$*Consistent*(s). By definition (498) and Rule $\neg\neg$E (78.2), it follows that $\exists p(s \models p \;\&\; s \models \neg p)$. Let q_1 be an arbitrary such proposition, so that we know:

(ϑ) $s \models q_1 \;\&\; s \models \neg q_1$

Then by (473.1), it follows that both $\Box s \models q_1$ and $\Box s \models \neg q_1$. But a conjunction of necessities is equivalent to a necessary conjunction (158.3), and so it follows that $\Box(s \models q_1 \;\&\; s \models \neg q_1)$. Hence, by $\exists$I, $\exists p \Box(s \models p \;\&\; s \models \neg p)$, and this conclusion remains once we discharge (ϑ) by $\exists$E. Thus, by the Buridan formula (168.1), it follows that $\Box\exists p(s \models p \;\&\; s \models \neg p)$. But we may apply the relevant instance of the modally-strict theorem that $\varphi \equiv \neg\neg\varphi$ and a Rule of Substitution (160.2) to obtain:

$$\Box\neg\neg\exists p(s \models p \;\&\; s \models \neg p)$$

Hence, by definition of *Consistent* (498) and a Rule of Substitution (160.3), it follows that $\Box\neg$*Consistent*(s). $(\leftarrow)$ Exercise. ⋈

(501.2) (Exercise)

(503.1) (Exercise)

(503.2) Assume $\exists p((s \models p) \;\&\; \neg\Diamond p)$. Suppose p_1 is such a proposition, so that we know (a) $s \models p_1$ and (b) $\neg\Diamond p_1$. By definitions (502) and (492), we have to show $\neg\Diamond\forall q(s \models q \rightarrow q)$. By (162.1), it suffices to show $\Box\neg\forall q(s \models q \rightarrow q)$. But since there is a modally strict proof of the equivalence of $\neg\forall q(s \models q \rightarrow q)$ and $\exists q(s \models q \,\&\, \neg q)$ (exercise), it suffices by the Rule of Substitution (160.2) to show $\Box\exists q(s \models q \,\&\, \neg q)$. Now from (a) we know $\Box s \models p_1$ by (473.1), and from (b) we know $\Box\neg p_1$ by (162.1). So by &I and (158.3) , we know $\Box(s \models p_1 \,\&\, \neg p_1)$. By $\exists$I, it follows that $\exists q\Box(s \models q \,\&\, \neg q)$. But by the Buridan formula (168.1), it follows that $\Box\exists q(s \models q \,\&\, \neg q)$. ⋈

(504.1) Assume *Possible*(s). Then by definition (502),

(ϑ) $\Diamond Actual(s)$

Note independently that (500) is a theorem, and since the free restricted variable is rigid, we can apply expanded RN (341.3.a) to obtain:

(ξ) $\Box(Actual(s) \rightarrow Consistent(s))$

Hence by an instance of the K$\Diamond$ schema (158.13), it follows from (ϑ) and (ξ) that $\Diamond Consistent(s)$. But then, by (501.2), *Consistent*(s). ⋈

(504.2) [The following proof rehearses some of the discussion in Remark (499).] Let q_1 be any proposition and let s_1 be the situation:

$$\imath s\forall p(s \models p \equiv p = (q_1 \,\&\, \neg q_1))$$

We leave it as an exercise to show s_1 is identical to a strictly canonical situation, i.e., to show, when φ is the formula $p = (q_1 \,\&\, \neg q_1)$, that there are modally strict proofs of $\forall p(\varphi \rightarrow \Box\varphi)$. Hence by (486.4), it follows by definition of s_1 that:

$$\forall p(s_1 \models p \equiv p = (q_1 \,\&\, \neg q_1))$$

So we know both that $s_1 \models (q_1 \,\&\, \neg q_1)$ and that s_1 encodes no other properties. Hence, we know both that $s_1 \models (q_1 \,\&\, \neg q_1)$ and that no other proposition is true in s_1. Consequently, there is no proposition p such that both p and $\neg p$ are true in s_1. So by definition (498), *Consistent*(s_1). It remains to show $\neg Possible(s_1)$. Assume, for reductio, that *Actual*(s_1). Then by definition of *Actual* (492), it follows that $q_1 \,\&\, \neg q_1$, which is a contradiction. Hence, $\neg Actual(s_1)$. Since this is a modally strict theorem, we may apply RN to obtain: $\Box\neg Actual(s_1)$, i.e., $\neg\Diamond Actual(s_1)$. So, by definition (502), $\neg Possible(s_1)$. ⋈

(505.1) – **(505.11)** (Exercises)

(506.4) Assume $s!s'$. Then by (506.1), $\exists p(s \models p \,\&\, s' \models \overline{p})$, say q_1, so that we know $s \models q_1 \,\&\, s' \models \overline{q_1}$. But $\neg\Diamond(q_1 \,\&\, \neg q_1)$. Since $\neg q_1 = \overline{q_1}$, $\neg\Diamond(q_1 \,\&\, \overline{q_1})$. Hence:

$$\neg\Diamond(q_1 \,\&\, \overline{q_1}) \,\&\, s \models q_1 \,\&\, s' \models \overline{q_1}$$

So by one application of $\exists$I:

$$\exists q(\neg\Diamond(q_1 \,\&\, q) \,\&\, s \models q_1 \,\&\, s' \models q)$$

And by a second application:

$$\exists p \exists q(\neg\Diamond(p \,\&\, q) \,\&\, s \models p \,\&\, s' \models q)$$

Hence, $s \boxtimes s'$, by (506.2). ⋈

(506.5) Assume $s \boxtimes s'$. Then by (506.2), $\exists p \exists q(\neg\Diamond(p \,\&\, q) \,\&\, s \models p \,\&\, s' \models q)$. So by modal negation, $\exists p \exists q(\Box\neg(p \,\&\, q) \,\&\, s \models p \,\&\, s' \models q)$. *A fortiori*, by the T schema, $\exists p \exists q(\neg(p \,\&\, q) \,\&\, s \models p \,\&\, s' \models q)$. Hence $s|s'$, by (506.3). ⋈

(506.6) By hypothetical syllogism, from (506.4) and (506.5). ⋈

(506.7) To construct witnesses to (.7), let:

- p_1 be some contingently true proposition (217.1),
- q_1 be some necessarily false proposition (208.2)
- $s_1 = \imath s \forall r(s \models r \equiv r = p_1)$, i.e., p_1 is the only proposition true in s_1, and
- $s_2 = \imath s \forall r(s \models r \equiv r = q_1)$, i.e., q_1 is the only proposition true in s_2.

Then since q_1 is necessarily false, so is $(p_1 \,\&\, q_1)$. So we know $\neg\Diamond(p_1 \,\&\, q_1)$, $s_1 \models p_1$, and $s_2 \models q_1$, from which it follows that $s_1 \boxtimes s_2$. But p_1 can't be the negation of q_1 – the negation of a necessary falsehood is a necessary truth and p_1 is, by hypothesis, contingently true. Nor can q_1 be the negation of p_1 – the negation of a contingent truth is a contingent falsehood and q_1 is, by hypothesis, a necessary falsehood. So we have $\neg\exists p(s \models p \,\&\, s' \models \overline{p})$, i.e., $\neg s!s'$. ⋈

(506.8) Use the same two witnesses constructed in the proof of (506.7): if q_1 is necessarily false, then $p_1 \,\&\, q_1$ is false, and so we have $\neg(p_1 \,\&\, q_1)$, $s_1 \models p_1$, and $s_2 \models q_1$, from which it follows that $s|s'$. Yet we still have $\neg s|s'$, by the same reasoning used above.

(506.9) To construct witnesses to (.9), we need to first find propositions p and q such that $\neg(p \,\&\, q)$ and $\Diamond(p \,\&\, q)$. Let p_2 be a contingently false proposition, and q_2 be a necessarily true proposition. So $\neg(p_2 \,\&\, q_2)$, since p_2 is contingently false, but $\Diamond(p_2 \,\&\, q_2)$, since $p_2 \,\&\, q_2$ is true at any world where p_2 is true. Then if $s_3 = \imath s \forall r(s \models r \equiv r = p_2)$ and $s_4 = \imath s \forall r(s \models r \equiv r = q_2)$, we have both:

- $s_3|s_4$. Since $\neg(p_2 \,\&\, q_2)$, $s_3 \models p_2$, and $s_4 \models q_2$, it follows that $\exists p \exists q(\neg(p \,\&\, q) \,\&\, s_3 \models p \,\&\, s_4 \models q)$.
- $\neg s_3 \boxtimes s_4$. Suppose, for reductio, $\exists p \exists q(\neg\Diamond(p \,\&\, q) \,\&\, s_3 \models p \,\&\, s_4 \models q)$. Then let p_3 and q_3 be such that $\neg\Diamond(p_3 \,\&\, q_3) \,\&\, s_3 \models p_3 \,\&\, s_4 \models q_3$. By the definition of s_3 and its strict canonicity, it follows that $\forall r(s_3 \models r \equiv r = p_2)$. Hence $p_3 = p_2$. And by the definition of s_4 and its strict canonicity, it follows that $\forall r(s_4 \models r \equiv r = q_2)$. Hence $q_3 = q_2$. But from the two identity claims just established and the fact that $\neg\Diamond(p_3 \,\&\, q_3)$, it follows that $\neg\Diamond(p_2 \,\&\, q_2)$, which contradicts the fact that $\Diamond(p_2 \,\&\, q_2)$. ⋈

(507.1) Assume $s!s'$. Then by (506.1), $\exists p(s \models p \;\&\; s' \models \overline{p})$, say q_1, so that we know:

(ϑ) $s \models q_1 \;\&\; s' \models \overline{q_1}$

Suppose, for reductio, $\exists s''(Consistent(s'') \;\&\; s \trianglelefteq s'' \;\&\; s' \trianglelefteq s'')$. Suppose s_1 is such a situation, so that we know:

$$Consistent(s_1) \;\&\; s \trianglelefteq s_1 \;\&\; s' \trianglelefteq s_1$$

So by definitions (498) and (475), we know:

(A) $\neg\exists p(s_1 \models p \;\&\; s_1 \models \neg p)$

(B) $\forall p(s \models p \rightarrow s_1 \models p)$

(C) $\forall p(s' \models p \rightarrow s_1 \models p)$

But (B) and the first conjunct of (ϑ) imply $s_1 \models q_1$. And (C) and the second conjunct of (ϑ) imply $s_1 \models \overline{q_1}$. The latter implies $s_1 \models \neg q_1$. Hence q_1 is a witness to that establishes $\exists p(s_1 \models p \;\&\; s_1 \models \overline{p})$, contradicting (A). $\bowtie$

(507.2) Assume $s \boxtimes s'$. Then by (506.1):

$$\exists p \exists q(\neg\Diamond(p \;\&\; q) \;\&\; s \models p \;\&\; s' \models q)$$

Suppose p_1 and q_1 are such propositions, so that we know:

(ϑ) $\neg\Diamond(p_1 \;\&\; q_1) \;\&\; s \models p_1 \;\&\; s' \models q_1$

Suppose, for reductio, $\exists s''(Possible(s'') \& s \trianglelefteq s'' \& s' \trianglelefteq s'')$. Let s_1 be such a situation, so that we know:

$$Possible(s_1) \;\&\; s \trianglelefteq s_1 \;\&\; s' \trianglelefteq s_1$$

So by definitions (502), (492), and (475), we know:

(A) $\Diamond\forall p(s_1 \models p \rightarrow p)$

(B) $\forall p(s \models p \rightarrow s_1 \models p)$

(C) $\forall p(s' \models p \rightarrow s_1 \models p)$

But (B) and the second conjunct of (ϑ) imply $s_1 \models p_1$. And (C) and the third conjunct of (ϑ) imply $s_1 \models q_1$. Since both $s_1 \models p_1$ and $s_1 \models \neg q_1$ are necessary, we can conjoin their necessitations and then, by the principle $(\Box\varphi \;\&\Box\psi) \rightarrow \Box(\varphi \;\&\; \psi)$, conclude:

(D) $\Box(s_1 \models p_1 \;\&\; s_1 \models q_1)$

Now independently, we can establish:

(E) $\Box(s_1 \models p_1 \,\&\, s_1 \models q_1) \rightarrow \Box(\forall p(s_1 \models p \rightarrow p) \rightarrow (p_1 \,\&\, q_1))$

Proof. Assume first $s_1 \models p_1 \,\&\, s_1 \models q_1$ and then assume $\forall p(s_1 \models p \rightarrow p)$. Our two assumptions imply $p_1 \,\&\, q_1$. So we may discharge our assumptions by two applications of conditional proof to obtain:

$$(s_1 \models p_1 \,\&\, s_1 \models q_1) \rightarrow (\forall p(s_1 \models p \rightarrow p) \rightarrow (p_1 \,\&\, q_1))$$

Since we established this by modally strict means and from no assumptions, it follows by RN that:

$$\Box((s_1 \models p_1 \,\&\, s_1 \models q_1) \rightarrow (\forall p(s_1 \models p \rightarrow p) \rightarrow (p_1 \,\&\, q_1)))$$

Hence, (E) follows by the K axiom.

So from (D) and (E), it follows that $\Box(\forall p(s_1 \models p \rightarrow p) \rightarrow (p_1 \,\&\, q_1))$. Hence by the K$\Diamond$ principle, it follows that $\Diamond\forall p(s_1 \models p \rightarrow p) \rightarrow \Diamond(p_1 \,\&\, q_1)$. But the antecedent of this last fact is just (A). Hence $\Diamond(p_1 \,\&\, q_1)$, which contradicts the first conjunct of (ϑ). ⋈

(507.3) By hypothetical syllogism, from (506.4) and (507.2). ⋈

(508.2) (Exercise)

(508.3) Assume $(s \models p \,\&\, \neg s \models \overline{p}) \vee (s \models \overline{p} \,\&\, \neg s \models p)$. Then we prove the consequent by cases. (1) Assume $s \models p$ and $\neg s \models \overline{p}$. Then by the latter and (508.2), $s^* \models p$. Since $s^* \models p$ and $s \models p$ are both true, $s^* \models p \equiv s \models p$. (2) Assume $s \models \overline{p}$ and $\neg s \models p$. Then by the former and (508.2), $\neg s^* \models p$. Since $s^* \models p$ and $s \models p$ are both false, we again have $s^* \models p \equiv s \models p$. ⋈

(508.4) We prove our theorem with the help of some preliminary facts. If we apply GEN to (508.2), we know:

(A) $\forall s\forall p(s^* \models p \equiv \neg s \models \overline{p})$

Moreover, since the universal quantifiers of (A) commute, it follows also that:

(B) $\forall p\forall s(s^* \models p \equiv \neg s \models \overline{p})$

With the help of (A) and (B), we next prove two lemmata:

Lemma 1: $\forall p(s^{**} \models p \equiv \neg s^* \models \overline{p})$
Proof. By instantiating (A) to s^*.

Lemma 2: $\forall p(s^* \models \overline{p} \equiv \neg s \models \overline{\overline{p}})$
Proof. If we instantiate (B) to $\overline{q}$, where q is arbitrary, we obtain:

$$s^* \models \overline{q} \equiv \neg s \models \overline{\overline{q}}$$

Since q was arbitrary, this last result implies:

$\forall q(s^* \models \overline{q} \equiv \neg s \models \overline{\overline{q}})$

But this is an alphabetic variant of what we're trying to prove.

Now we may prove both directions of our theorem as follows: ($\rightarrow$) Assume s is classical w.r.t. double negation, i.e.,

(ϑ) $\forall p(s \models p \equiv s \models \overline{\overline{p}})$

From (ϑ), Lemma 1, and Lemma2, we can establish $s^{**} = s$ by appealing to (474). So we have to show, for an arbitrary q, that $s^{**} \models q \equiv s \models q$:

$$\begin{array}{lcll} s^{**} \models q & \equiv & \neg s^* \models \overline{q} & \text{by Lemma 1} \\ & \equiv & \neg\neg s \models \overline{\overline{q}} & \text{by Lemma 2} \\ & \equiv & s \models \overline{\overline{q}} & \text{by (88.3.b)} \\ & \equiv & s \models q & \text{by } (\vartheta) \end{array}$$

($\leftarrow$) Assume $s^{**} = s$. To show $\forall p(s \models p \equiv s \models \overline{\overline{p}})$, we may reason with respect to an arbitrary q, as follows:

$$\begin{array}{lcll} s \models q & \equiv & s^{**} \models q & \text{from our assumption, by Rule =E} \\ & \equiv & \neg s^* \models \overline{q} & \text{by Lemma 1} \\ & \equiv & \neg\neg s \models \overline{\overline{q}} & \text{by Lemma 2} \\ & \equiv & s \models \overline{\overline{q}} & \text{by (88.3.b)} \end{array} \qquad \bowtie$$

(508.5) Assume $\forall p(s \models p \equiv s \models \overline{\overline{p}})$ and that s has a gap w.r.t. q, i.e., $\neg s \models q$ and $\neg s \models \overline{q}$. From the latter and (508.2), it follows that $s^* \models q$. From the former and the assumption that $\forall p(s \models p \equiv s \models \overline{\overline{p}})$, it follows that $\neg s \models \overline{\overline{q}}$. But by (508.2), we know $\forall p(s^* \models \overline{p} \equiv \neg s \models \overline{\overline{p}})$ – see the proof of Lemma 2 in the previous theorem. Hence $s^* \models \overline{q}$. So s^* has a glut w.r.t. q. $\bowtie$

(508.6) Assume $\forall p(s \models p \equiv s \models \overline{\overline{p}})$ and that s has a glut w.r.t. q, i.e., that $s \models q$ and $s \models \overline{q}$. Then to show s^* has a gap w.r.t. q, we show both $\neg s^* \models q$ and $\neg s^* \models \overline{q}$. Now (508.2) tells us that $\forall p(s^* \models p \equiv \neg s \models \overline{p})$. So from the assumption that $s \models \overline{q}$, it follows that $\neg s^* \models q$. So it remains to show $\neg s^* \models \overline{q}$. From the assumption that $\forall p(s \models p \equiv s \models \overline{\overline{p}})$ and the assumption that $s \models q$, it follows that $s \models \overline{\overline{q}}$. However, if we instantiate (508.2) to $\overline{q}$, then it follows that $s^* \models \overline{q} \equiv \neg s \models \overline{\overline{q}}$. Hence $\neg s^* \models \overline{q}$. $\bowtie$

(508.7) Assume s is consistent w.r.t. double negation, and suppose, for reductio, that $s!s^*$. Then by definition (506.1), $\exists p(s \models p \;\&\; s^* \models \overline{p})$. Suppose q_1 is such a proposition, so that we know $s \models q_1$ and $s^* \models \overline{q_1}$. By the principal fact about s^* (508.2), the latter implies $\neg s \models \overline{\overline{q_1}}$. But since s is consistent w.r.t. double negation, it follows that $\neg s \models q_1$. Contradiction. $\bowtie$

(508.8) Assume:

(ϑ) $\neg s!s'$

(ξ) $\forall p(s' \models p \equiv s' \models \overline{\overline{p}})$

To show $s' \trianglelefteq s^*$, i.e., $\forall q(s' \models q \rightarrow s^* \models q)$, it suffices by GEN to show $s' \models q \rightarrow s^* \models q$. So further assume:

(ζ) $s' \models q$

Now it follows from (ϑ) by definition (506.1) that $\neg\exists p(s \models p \;\&\; s' \models \overline{p})$, i.e., $\forall p\neg(s \models p \;\&\; s' \models \overline{p})$. If we instantiate this last fact to $\overline{q}$, then we know $\neg(s \models \overline{q} \;\&\; s' \models \overline{\overline{q}})$, i.e.,

(ω) $\neg s \models \overline{q} \vee \neg s' \models \overline{\overline{q}}$

Now suppose, for reductio, $\neg s^* \models q$. Then by (508.2), $\neg\neg s \models \overline{q}$. From this and ($\omega$), it follows that $\neg s' \models \overline{\overline{q}}$. But ($\xi$) and ($\zeta$) jointly imply $s' \models \overline{\overline{q}}$. Contradiction. ⋈

(511.1) Assume $\forall p(s \models p \equiv p)$. By GEN, it suffices to show $(s \models \neg q) \equiv \neg s \models q$. Now the following are instances of our assumption:

(ϑ) $s \models q \equiv q$

(ξ) $s \models \neg q \equiv \neg q$

It follows from (ϑ) by (88.4.b) that $\neg s \models q \equiv \neg q$, i.e., $\neg q \equiv \neg s \models q$. Hence from ($\xi$) and this conclusion it follows that $(s \models \neg q) \equiv \neg s \models q$. ⋈

(511.2) Assume $\forall p(s \models p \equiv p)$. Now the following are instances of our assumption:

(ϑ) $s \models q \equiv q$

(ξ) $s \models r \equiv r$

(ζ) $(s \models (q \rightarrow r)) \equiv (q \rightarrow r)$

By two applications of GEN, it suffices to show $s \models (q \rightarrow r) \equiv ((s \models q) \rightarrow (s \models r))$. ($\rightarrow$) Assume both $s \models (q \rightarrow r)$ and $s \models q$, to show $s \models r$. By the first and (ζ), $q \rightarrow r$. By the second and (ϑ), q. Hence, r. But from r and (ξ), it follows that $s \models r$. ($\leftarrow$) Assume $(s \models q) \rightarrow (s \models r)$. For reductio, assume $\neg s \models (q \rightarrow r)$. Since we're under the global assumption $\forall p(s \models p \equiv p)$, it follows from our reductio assumption that $s \models \neg(q \rightarrow r)$, by (511.1). But this implies, by our global assumption and the fact that $(\neg(q \rightarrow r))\downarrow$ (104.2), that $\neg(q \rightarrow r)$. So we know both q and $\neg r$. By (ϑ), the former implies $s \models q$, and this implies, by our local assumption, that $s \models r$. Hence, by (ξ), r. Contradiction. ⋈

(511.3) Assume $\forall p(s \models p \equiv p)$. Since both $\varphi\downarrow$ and $(\forall\alpha\varphi)\downarrow$ (104.2), we may instantiate them for p in our assumption to obtain:

(ϑ) $s \models \varphi \equiv \varphi$

(ξ) $s \models \forall\alpha\varphi \equiv \forall\alpha\varphi$

($\rightarrow$) Assume $s \models \forall\alpha\varphi$. But this implies, by ($\xi$), that $\forall\alpha\varphi$. Hence φ, by the Variant of $\forall$E (93.3). So, by (ϑ), $s \models \varphi$. Since α is not free in any assumption, it follows by GEN that $\forall\alpha(s \models \varphi)$.

($\leftarrow$) Assume $\forall\alpha(s \models \varphi)$. Then by the Variant of $\forall$E (93.3), $s \models \varphi$. So by (ϑ), φ. Since α is not free in any assumption, it follows by GEN that $\forall\alpha\varphi$. Hence, by (ξ), $s \models \forall\alpha\varphi$. ⋈

(511.4) Assume $\forall p(s \models p \equiv p)$. By GEN, we have to show $(s \models \Box q) \rightarrow (\Box s \models q)$. Note that the following are instances of our assumption:

(ϑ) $s \models q \equiv q$

(ξ) $s \models \Box q \equiv \Box q$

To complete our proof, assume $s \models \Box q$. Then by (ξ), $\Box q$, and by the T schema, q. So by (ϑ), $s \models q$. Since *truth-in-s* is rigid (473.1), $\Box s \models q$. ⋈

(511.5) Assume $\forall p(s \models p \equiv p)$. Then since we know there are contingently true propositions (217.1), let r be an arbitrary such proposition, so that we know by (213.1) that:

(A) r

(B) $\Diamond\neg r$, i.e., $\neg\Box r$

Since $r\downarrow$, (A) and our assumption imply $s \models r$. So by the rigidity of *truth-in-s* (473.1):

(C) $\Box s \models r$

But, clearly, $(\neg\Box r)\downarrow$, and so (B) and our assumption imply $s \models \neg\Box r$. Hence, by (511.1), $\neg s \models \Box r$. So by (C) and this last result: $\exists q(\Box(s \models q)\ \&\ \neg(s \models \Box q))$. ⋈

(511.6) This is an instance of theorem (486.1), when φ is the variable p. ⋈

(512.2) To find our witness to this existential claim, note that by theorem (511.6), we know $\exists s\forall p(s \models p \equiv p)$. By eliminating the restricted variables, this becomes:

$$\exists x(Situation(x)\ \&\ \forall p(x \models p \equiv p))$$

Let a be such an object, so that we know:

(ϑ) $Situation(a)\ \&\ \forall p(a \models p \equiv p)$

But by the T$\Diamond$ schema (163.1), the right conjuct of (ϑ) implies $\Diamond\forall p(a \models p \equiv p)$. Hence, we've established:

$$Situation(a) \,\&\, \Diamond\forall p(a \models p \equiv p)$$

and so by definition (512), it follows that a is a possible world. So by now familiar reasoning, $\exists x PossibleWorld(x)$. ⋈

(512.3) (Exercise)

(513.1) We prove only the left-to-right direction since the right-to-left direction is just an instance of the T schema. ($\rightarrow$) Assume $PossibleWorld(x)$. By definition (512), we know:

(ϑ) $Situation(x) \,\&\, \Diamond\forall p(x \models p \equiv p)$

By (469.1), the first conjunct of (ϑ) implies:

(ξ) $\Box Situation(x)$

Moreover, by the 5 schema (45.3), the second conjunct of (ϑ) implies:

(ζ) $\Box\Diamond\forall p(x \models p \equiv p)$

Hence, by the right-to-left direction of (158.3), the conjunction of (ξ) and (ζ) implies:

$$\Box(Situation(x) \,\&\, \Diamond\forall p(x \models p \equiv p))$$

So by the definition of possible world (512) and the Rule of Substitution for Defined Subformulas (160.3), $\Box PossibleWorld(x)$. ⋈

(513.2) ($\rightarrow$) It follows *a fortiori* from (513.1) that:

$$PossibleWorld(x) \rightarrow \Box PossibleWorld(x)$$

Since this result is a modally strict theorem, it follows by Rule (166.2) that $\Diamond PossibleWorld(x) \rightarrow PossibleWorld(x)$. ($\leftarrow$) (Exercise) ⋈

(513.3) (Exercise)

(513.4) ($\rightarrow$) Assume $\mathcal{A}PossibleWorld(x)$. Independently, it is a consequence of (513.1) that $PossibleWorld(x) \rightarrow \Box PossibleWorld(x)$. So by RA:

$$\mathcal{A}(PossibleWorld(x) \rightarrow \Box PossibleWorld(x))$$

So by theorem (131):

$$\mathcal{A}PossibleWorld(x) \rightarrow \mathcal{A}\Box PossibleWorld(x)$$

Since the antecedent holds by assumption, it follows that $\mathcal{A}\Box PossibleWorld(x)$. So by (46.2), $\Box PossibleWorld(x)$. Hence by the T schema, $PossibleWorld(x)$. ($\leftarrow$) Assume $PossibleWorld(x)$. Then by (513.1), $\Box PossibleWorld(x)$. So by (132), $\mathcal{A}PossibleWorld(x)$. ⋈

(516) (Exercise)

(517) By the conventions for rigid restricted variables (341.2) and the fact that w is a rigid restricted variable for the rigid restriction condition $PossibleWorld(x)$, it is a necessary axiom that $PossibleWorld(w)$. So by definition (512), that $\Diamond\forall p(w\models p \equiv p)$. *A fortiori* (exercise), $\Diamond\forall p(w\models p \rightarrow p)$. So by the definition of *Actual* (492), $\Diamond Actual(w)$. By the definition of *Possible* (502), $Possible(w)$. ⋈

(518.1) By the previous theorem (517), $Possible(w)$. So by (504.1), $Consistent(w)$. ⋈

(518.2) (Exercise)

(519.1) – **(519.5)** (Exercises) [Cf. (473.1) – (473.5)]

(521) Since w is a possible world, it follows by definition (512) that:

(ϑ) $\Diamond\forall p(w\models p \equiv p)$

To show $Maximal(w)$ (520), we have to show $\forall q(w\models q \vee w\models\neg q)$. So, by GEN, it suffices to show, $w\models q \vee w\models\neg q$. Our proof strategy will be to:

(a) show that $\Diamond(w\models q \vee w\models\neg q)$, and then

(b) appeal to various modal facts, including the rigidity of *truth at* (519.2), to derive that $w\models q \vee w\models\neg q$ from (a).

For (a), our proof strategy is to:

(i) show $\Box(\varphi \rightarrow \psi)$, where φ is $\forall p(w\models p \equiv p)$ and ψ is $w\models q \vee w\models\neg q$, and

(ii) use the modal law $\Box(\varphi \rightarrow \psi) \rightarrow (\Diamond\varphi \rightarrow \Diamond\psi)$ (158.13) to conclude $\Diamond\psi$ from (i) and (ϑ), i.e., from $\Box(\varphi \rightarrow \psi)$ and $\Diamond\varphi$.

For (i), assume $\forall p(w\models p \equiv p)$. Then both $w\models q \equiv q$ and $w\models\neg q \equiv \neg q$. Since $q\vee\neg q$, it follows by disjunctive syllogism (89.1) that $w\models q \vee w\models\neg q$. Since we've now established, by conditional proof, that $\forall p(w\models p \equiv p) \rightarrow (w\models q \vee w\models\neg q)$ is a modally strict theorem, it follows by the fact that the free restricted variables are rigid and expanded RN (341.3.a) that:

(ξ) $\Box(\forall p(w\models p \equiv p) \rightarrow (w\models q \vee w\models\neg q))$

Now for (ii), it follows from (ξ) and (ϑ) by the modal law (158.13) that:

(ζ) $\Diamond(w\models q \vee w\models\neg q)$

Now for (b), if we apply to (ζ) the modal law (162.2), which asserts that possibility distributes over a disjunction, it follows that:

$$\Diamond w \models q \lor \Diamond w \models \neg q$$

But by (519.2), the left disjunct is equivalent to $w \models q$ and the right is equivalent to $w \models \neg q$. So, by disjunctive syllogism (89.1), it follows that $w \models q \lor w \models \neg q$. ⋈

(522.1) Assume *Maximal*(s). Then by definition of *Maximal* (520):

(ϑ) $\forall p(s \models p \lor s \models \neg p)$

We want to show $\Box\forall p(s \models p \lor s \models \neg p)$. By the Barcan Formula (167.1), it suffices to show $\forall p\Box(s \models p \lor s \models \neg p)$. By GEN, it suffices to show $\Box(s \models p \lor s \models \neg p)$. Now from ($\vartheta$), it follows by $\forall$E that $s \models p \lor s \models \neg p$. But by (473.1), the first disjunct implies $\Box s \models p$ and the second disjunct implies $\Box s \models \neg p$. Hence, by disjunctive syllogism, $\Box s \models p \lor \Box s \models \neg p$. So by (158.15), $\Box(s \models p \lor s \models \neg p)$. ⋈

(522.2) ($\rightarrow$) By (521) and (517). ($\leftarrow$) Before we begin, we establish a few facts needed for the proof, the first of which is:

$$Maximal(s) \rightarrow (\forall p(s \models p \rightarrow p) \rightarrow \forall p(s \models p \equiv p))$$

Proof. Assume (a) *Maximal*(s) and (b) $\forall p(s \models p \rightarrow p)$. To show $\forall p(s \models p \equiv p)$ it suffices, by GEN, to show $s \models p \equiv p$. ($\rightarrow$) This direction is immediate by instantiating (b) to p. ($\leftarrow$) Assume p. Now assume, for reductio, that $\neg s \models p$. Then by (a) and the definition of maximality (520), it follows that $s \models \neg p$. But we know, as an instance of (b) that $s \models \neg p \rightarrow \neg p$. Hence, $\neg p$. Contradiction.

Since this first fact is a modally strict theorem and the free restricted variables are rigid, it follows by expanded RN (341.3.a) that:

$$\Box[Maximal(s) \rightarrow (\forall p(s \models p \rightarrow p) \rightarrow \forall p(s \models p \equiv p))]$$

So by the K axiom, we know:

(ϑ) $\Box Maximal(s) \rightarrow \Box(\forall p(s \models p \rightarrow p) \rightarrow \forall p(s \models p \equiv p))$

Now to establish our theorem, assume *Maximal*(s) & *Possible*(s). Then from *Maximal*(s) and the previous theorem (522.1), it follows that $\Box$*Maximal*(s). From this and (ϑ) it follows that:

$$\Box(\forall p(s \models p \rightarrow p) \rightarrow \forall p(s \models p \equiv p))$$

From this, it follows by theorem (158.13) that:

(ξ) $\Diamond\forall p(s \models p \rightarrow p) \rightarrow \Diamond\forall p(s \models p \equiv p)$

But *Possible*(s) by assumption, from which it follows by definition of *Possible* that $\Diamond Actual(s)$, and by definition of *Actual* that $\Diamond\forall p(s \models p \rightarrow p)$ But from this last fact and (ξ), it follows that $\Diamond\forall p(s \models p \equiv p)$, i.e., *PossibleWorld*$(s)$. ⋈

(525) (Exercise)

(526.1) We prove that our theorem holds for any choice of $n \geq 1$. For conditional proof, assume both:

(ϑ) $s \models p_1 \;\&\; \ldots \;\&\; s \models p_n \;\&\; ((p_1 \;\&\; \ldots \;\&\; p_n) \rightarrow q)$

(ξ) $\forall p(s \models p \equiv p)$

Then instantiate each of $p_1, \ldots, p_n$ into (ξ), to obtain:

$$\begin{array}{l} s \models p_1 \equiv p_1 \\ \quad\vdots \\ s \models p_n \equiv p_n \end{array}$$

But each left-side condition in the above biconditionals is one of the first n conjuncts in (ϑ), and so it follows that $p_1 \;\&\; \ldots \;\&\; p_n$. Hence, by the last conjunct in (ϑ), it follows that q. But now instantiate q into (ξ), and we obtain $s \models q \equiv q$. Hence $s \models q$. ⋈

(526.2) We prove that our theorem holds for any choice of $n \geq 1$. Since the free restricted variables in (526.1) are rigid, we can apply expanded RN (341.3.a) to obtain:

(ϑ) $\Box[(s \models p_1 \;\&\; \ldots \;\&\; s \models p_n \;\&\; ((p_1 \;\&\; \ldots \;\&\; p_n) \rightarrow q)) \rightarrow (\forall p(s \models p \equiv p) \rightarrow s \models q)]$

If we:

let φ_1 be the formula $s \models p_1$
...
let φ_n be the formula $s \models p_n$
let ψ be the formula $(p_1 \;\&\; \ldots \;\&\; p_n) \rightarrow q$
let χ be the formula $\forall p(s \models p \equiv p)$
let θ be the formula $s \models q$

then (ϑ) has the form:

$$\Box((\varphi_1 \;\&\; \ldots \;\&\; \varphi_n \;\&\; \psi) \rightarrow (\chi \rightarrow \theta))$$

By the K axiom (45.1), it follows that:

$$\Box(\varphi_1 \;\&\; \ldots \;\&\; \varphi_n \;\&\; \psi) \rightarrow \Box(\chi \rightarrow \theta)$$

But since it is a modally strict fact that a necessary conjunction is equivalent to a conjunction of necessary truths (158.3), we may use the Rule of Substitution (160.2) to infer:

(ξ) $(\Box\varphi_1 \;\&\; \ldots \;\&\; \Box\varphi_n \;\&\; \Box\psi) \rightarrow \Box(\chi \rightarrow \theta)$

But if we replace the metavariables in (ξ) with the formulas that they represent, we obtain:

(ζ) $(\Box s \models p_1 \;\&\; \ldots \;\&\; \Box s \models p_n \;\&\; \Box((p_1 \;\&\; \ldots \;\&\; p_n) \rightarrow q)) \rightarrow$
$\Box(\forall p(s \models p \equiv p) \rightarrow s \models q)$

But by definition (524.1), we know that it is a modally strict theorem that $\Box((p_1 \;\&\; \ldots \;\&\; p_n) \rightarrow q)$ is equivalent to $(p_1 \;\&\; \ldots \;\&\; p_n) \Rightarrow q$. So by a Rule of Substitution, it follows from (ζ) that:

$$(\Box s \models p_1 \;\&\; \ldots \;\&\; \Box s \models p_n \;\&\; ((p_1 \;\&\; \ldots \;\&\; p_n) \Rightarrow q)) \rightarrow \Box(\forall p(s \models p \equiv p) \rightarrow s \models q)$$ ⋈

(528) By the definition of *PossibleWorld*, we know:

(ϑ) $\Diamond\forall p(w \models p \equiv p)$

To show that *n-ModallyClosed*(w), fix $n \geq 1$ and assume:

(ξ) $w \models p_1 \;\&\; \ldots \;\&\; w \models p_n \;\&\; ((p_1 \;\&\; \ldots \;\&\; p_n) \Rightarrow q)$

By (519.1), we can infer the necessitation of each of the first n conjuncts of (ξ), and so by &I, we have:

$$\Box w \models p_1 \;\&\; \ldots \;\&\; \Box w \models p_n$$

Conjoining this result with the last conjunct of (ξ), we have:

$$\Box w \models p_1 \;\&\; \ldots \;\&\; \Box w \models p_n \;\&\; ((p_1 \;\&\; \ldots \;\&\; p_n) \Rightarrow q)$$

Hence, by (526.2), it follows that:

(ζ) $\Box(\forall p(w \models p \equiv p) \rightarrow w \models q)$

From (ζ) and (ϑ), it follows that: $\Diamond w \models q$, by (158.13). But by (519.2), it follows that $w \models q$. So we have established, by conditional proof:

$$(w \models p_1 \;\&\; \ldots \;\&\; w \models p_n \;\&\; ((p_1 \;\&\; \ldots \;\&\; p_n) \Rightarrow q)) \rightarrow w \models q$$

This conclusion rests on no assumptions and so the variables $p_1, \ldots, p_n$ and q do not occur free in any assumption. So our theorem follows by $n+1$ applications of GEN. ⋈

(529.1) We've already established *Maximal*(w) and *Consistent*(w), by (521) and (518.1), respectively. So, by definitions (520) and (498), we know, respectively, the following:

(ϑ) $\forall q(w \models q \vee w \models \neg q)$

(ζ) $\neg\exists q(w \models q \;\&\; w \models \neg q)$

Now to prove our theorem, we prove both directions:

($\rightarrow$) Assume $w \models \neg p$. For reductio, suppose $w \models p$. Then $w \models p \;\&\; w \models \neg p$, and so $\exists q(w \models q \;\&\; w \models \neg q)$, contradicting ($\zeta$).

($\leftarrow$) Assume $\neg w \models p$. Then by (ϑ), $w \models \neg p$. ⋈

(**529**.2) By applying GEN to (529.1), we know:

(ϑ) $\forall p(w \models \neg p \equiv \neg w \models p)$

($\rightarrow$) Assume $w \models p$. For reductio, assume $\neg\neg w \models \neg p$. Then $w \models \neg p$. By (ϑ), $\neg w \models p$. Contradiction. ($\leftarrow$) Assume $\neg w \models \neg p$. Then by (ϑ), $\neg\neg w \models p$. Hence $w \models p$. ⋈

(**531.1**) Recall that in the proof of (512.2), we established that there was an object, arbitrarily named a, that is a possible world. But as part of the proof, we established (ϑ) that $\forall p(a \models p \equiv p)$. So some possible world makes true all and only the truths. ⋈

(**531**.2) By the definition of uniqueness, we have to show:

$\exists w(Actual(w) \;\&\; \forall w'(Actual(w') \rightarrow w' = w))$

Now to find our witness to this existential claim, note that by theorem (531.1), we know $\exists w \forall p(s \models p \equiv p)$. Let w_0 be such a world, so that we know:

(ϑ) $\forall p(w_0 \models p \equiv p)$

By &I and $\exists$I, it suffices to show:

(A) $Actual(w_0)$

(B) $\forall w(Actual(w) \rightarrow w = w_0)$

To show (A), we have to show, by (492), that $\forall p(w_0 \models p \rightarrow p)$. But this follows *a fortiori* from (ϑ).

To show (B), it suffices to show $Actual(w) \rightarrow w = w_0$, by GEN. Assume $Actual(w)$ and suppose, for reductio, $w \neq w_0$. Since w and w_0 are distinct possible worlds, it follows by (516) that there is a proposition, say r, true in one but not in the other. Without loss of generality, assume that $w \models r$ and $\neg w_0 \models r$. Since we know by (A) above that w_0 is a possible world, it follows by theorem (521) that it is maximal. So $w_0 \models \neg r$. But w is actual by hypothesis and w_0 is actual by (A) above. Hence, by definition of *Actual* (492), r and $\neg r$ are both true. Contradiction. ⋈

(**533**) By applying the Rule of Actualization to theorem (531.2), we may conclude:

$$\mathcal{A}\exists! w Actual(w)$$

So by (176.2), it follows that:

$$\iota w Actual(w)\downarrow \qquad \bowtie$$

(535) By definition (302.1), we know:[466]

$$\top = \iota x(A!x \,\&\, \forall F(xF \equiv \exists q(q \,\&\, F = [\lambda y\, q])))$$

Here we use an alphabetic variant of the definition of $\top$, to avoid a clash of variables later. Independently, by expanding the definition (534) to eliminate the restricted variable, we know:

$$\boldsymbol{w}_\alpha = \iota x(PossibleWorld(x) \,\&\, Actual(w))$$

Let φ and ψ, respectively, be the matrices of the descriptions appearing on the right side of the foregoing identities, so that they assert $\top = \iota x\varphi$ and $\boldsymbol{w}_\alpha = \iota x\psi$, respectively. So to show $\top = \boldsymbol{w}_\alpha$, it suffices to show $\iota x\varphi = \iota x\psi$. We'll prove this by an appeal to (149.3); since we know $\iota x\varphi\downarrow$, (149.3) implies that it suffices to show $\Box\forall x(\varphi \equiv \psi)$. And by GEN and RN, it suffices to show $\varphi \equiv \psi$.

($\rightarrow$) Assume φ, i.e., $A!x$ and

(ϑ) $\forall F(xF \equiv \exists q(q \,\&\, F = [\lambda y\, q]))$

To show ψ, we have to show both:

(A) $PossibleWorld(x)$

(B) $Actual(x)$

(A) By definition (512), we have to show $Situation(x)$ and $\Diamond\forall p(x \models p \equiv p)$. For the former, we already know $A!x$, and so by definition (467), we have to show $\forall F(xF \rightarrow Propositional(F))$. So, by GEN, assume xF. Then $\exists q(q \,\&\, F = [\lambda y\, q])$, by ($\vartheta$). *A fortiori*, $\exists q(F = [\lambda y\, q])$. Hence $Propositional(F)$, by definition (275). So it remains to show $\Diamond\forall p(x \models p \equiv p)$. By the T$\Diamond$ schema, it suffices to show $\forall p(x \models p \equiv p)$, and by GEN, to show $x \models p \equiv p$. ($\rightarrow$) Assume $x \models p$. Since x is a situation, we know by (471) that $x[\lambda y\, p]$. Note that we can instantiate $[\lambda y\, p]$ into (ϑ); the quantifier $\exists q$ won't capture the free variable p. So by (ϑ), $\exists q(q \,\&\, [\lambda y\, p] = [\lambda y\, q])$. Suppose q_1 is such a proposition, so that we know q_1 and $[\lambda y\, p] = [\lambda y\, q_1]$. Then by definition of proposition identity, $p = q_1$. Hence p. ($\leftarrow$) Assume p. Then $p \,\&\, [\lambda y\, p] = [\lambda y\, p]$. So $\exists q(q \,\&\, [\lambda y\, p] = [\lambda y\, q])$. Hence,

[466] I'm indebted here to Daniel West. West showed that this theorem need not be labeled as a ★-theorem (as I had done in an earlier draft) and sent a modally strict proof as evidence. I've adopted West's proof strategy in what follows, both in outline and in many of the details.

by (ϑ), $x[\lambda y\, p]$. And since x is a situation, it follows that $x \models p$, by now familiar reasoning.

(B) By definition (492), we have to show $\forall p(x \models p \rightarrow p)$. But this was proved as part of the reasoning used to establish (A).

($\leftarrow$) Assume ψ, i.e., $PossibleWorld(x) \,\&\, Actual(w)$. To show φ, we have to show both $A!x$ and $\forall F(xF \equiv \exists q(q \,\&\, F = [\lambda y\, q]))$. $A!x$ follows from the facts that $PossibleWorld(x) \rightarrow Situation(x)$, by definition (512), and that $Situation(x) \rightarrow A!x$, by definition (467). So, by GEN, it remains to show $xF \equiv \exists q(q \,\&\, F = [\lambda y\, q])$. ($\rightarrow$) Assume xF, to find a witness to $\exists q(q \,\&\, F = [\lambda y\, q])$. But since we've already established that x is a situation, it also follows by definition (467) that $\forall F(xF \rightarrow \exists p(F = [\lambda y\, p]))$. Since xF by assumption, it follows that $\exists p(F = [\lambda y\, p])$. Let q_2 be such a proposition, so that we know $F = [\lambda y\, q_2]$. Then all we have to do to show that q_2 is the desired witness is to show that q_2 is true. But $x[\lambda y\, q_2]$ follows from what we've assumed and established thus far, and since x is a situation, we therefore know $x \models q_2$ by now familiar reasoning. But since x is actual, it follows by definition (492) that q_2 is true. ($\leftarrow$) Assume $\exists q(q \,\&\, F = [\lambda y\, q])$. Then, *a fortiori*, $\exists q(F = [\lambda y\, q])$. So by ($\vartheta$), xF. ⋈

(536.1)★ (Exercise)

(536.2)★ By (305.3)★, we know $p \equiv \top\Sigma p$. But by a previous theorem (535), we know $\top = \boldsymbol{w}_\alpha$. Hence, $p \equiv \boldsymbol{w}_\alpha \Sigma p$. But since $\boldsymbol{w}_\alpha$ is a situation, we know $\boldsymbol{w}_\alpha \Sigma p \equiv \boldsymbol{w}_\alpha \models p$, by definition (470) and Rule ≡S. Hence, $p \equiv \boldsymbol{w}_\alpha \models p$. ⋈

(537.1) By definition (534) and the method of eliminating restricted w variables noted in Remark (514), we know:

$$\boldsymbol{w}_\alpha = \imath x(PossibleWorld(x) \,\&\, Actual(x))$$

Since $\boldsymbol{w}_\alpha\downarrow$, it follows from the relevant instance of (152.3) that:

$$\mathcal{A}(PossibleWorld(\boldsymbol{w}_\alpha) \,\&\, Actual(\boldsymbol{w}_\alpha))$$

By (139.2), it follows that:

$$\mathcal{A}PossibleWorld(\boldsymbol{w}_\alpha) \,\&\, \mathcal{A}Actual(\boldsymbol{w}_\alpha))$$

So by (513.4), the first conjunct implies $PossibleWorld(\boldsymbol{w}_\alpha)$. ⋈

(537.2) By (537.1) we know $PossibleWorld(\boldsymbol{w}_\alpha)$. Hence by (521), $Maximal(\boldsymbol{w}_\alpha)$. ⋈

(538) By the first Exercise at the end of (305)★, we know there is a *modally strict* proof of $\mathcal{A}p \equiv \top\Sigma p$. But by (535), we established that $\top = \boldsymbol{w}_\alpha$. So $\mathcal{A}p \equiv \boldsymbol{w}_\alpha \Sigma p$. But $\boldsymbol{w}_\alpha$ is a situation and so it follows from (471) that $\boldsymbol{w}_\alpha \Sigma p \equiv \boldsymbol{w}_\alpha \models p$. So by biconditional syllogism, $\mathcal{A}p \equiv \boldsymbol{w}_\alpha \models p$. ⋈

(539)★ Assume $w \neq \boldsymbol{w}_\alpha$ but suppose, for reductio, that $Actual(w)$. Since w and $\boldsymbol{w}_\alpha$ are distinct possible worlds, it follows by (516) that there is a proposition q true in one but not in the other. Without loss of generality, suppose $w \models q$ and $\neg \boldsymbol{w}_\alpha \models q$. From the actuality of w and $w \models q$, it follows that q. But by (536.2)★ and $\neg \boldsymbol{w}_\alpha \models q$, it follows that $\neg q$. Contradiction. ⋈

(540)★ ($\rightarrow$) Assume $Actual(s)$. To show $s \trianglelefteq \boldsymbol{w}_\alpha$, it suffices, by definition (475) and GEN, to show: $s \models p \rightarrow \boldsymbol{w}_\alpha \models p$. So assume $s \models p$. Since s is actual, p is true. But by (536.2)★, all and only true propositions are true at $\boldsymbol{w}_\alpha$. Hence, $\boldsymbol{w}_\alpha \models p$. ($\leftarrow$) Assume $s \trianglelefteq \boldsymbol{w}_\alpha$. To show $Actual(s)$, it suffices by definition (492) and GEN to show: $s \models p \rightarrow p$. So assume $s \models p$. Then it follows from $s \trianglelefteq \boldsymbol{w}_\alpha$ and the definition of $\trianglelefteq$ (475) that $\boldsymbol{w}_\alpha \models p$. But we know by (536.2)★ that all and only true propositions are true in $\boldsymbol{w}_\alpha$. Hence, p is true. ⋈

(541.1)★ ($\rightarrow$) By (536.2)★, we know $p \equiv \boldsymbol{w}_\alpha \models p$. But since $[\lambda y\, p]\downarrow$, we know by β-Conversion that $[\lambda y\, p]\boldsymbol{w}_\alpha \equiv p$. So by biconditional syllogism, $[\lambda y\, p]\boldsymbol{w}_\alpha \equiv \boldsymbol{w}_\alpha \models p$. Our theorem is the commuted form. ⋈

(541.2)★ ($\rightarrow$) Suppose p. Then since $[\lambda y\, p]\downarrow$, Rule $\overleftarrow{\beta}$C implies $[\lambda y\, p]\boldsymbol{w}_\alpha$. But note that $[\lambda y\, p]\boldsymbol{w}_\alpha$ is a 0-ary relation term and so may be instantiated into the universal generalization of (536.2)★ to obtain:

$$[\lambda y\, p]\boldsymbol{w}_\alpha \equiv \boldsymbol{w}_\alpha \models [\lambda y\, p]\boldsymbol{w}_\alpha$$

Hence, $\boldsymbol{w}_\alpha \models [\lambda y\, p]\boldsymbol{w}_\alpha$. ($\leftarrow$) By reverse reasoning. ⋈

(542.1) As an instance of theorem (158.13), we know:

$$\Box(p \rightarrow \exists w(w \models p)) \rightarrow (\Diamond p \rightarrow \Diamond \exists w(w \models p))$$

So to show the consequent $\Diamond p \rightarrow \Diamond \exists w(w \models p)$, it suffices to show the antecedent $\Box(p \rightarrow \exists w(w \models p))$. But by RN, it suffices to give a modally strict proof of $p \rightarrow \exists w(w \models p)$. So assume p and, for reductio, $\neg \exists w(w \models p)$, i.e., $\forall w \neg(w \models p)$. Now we know, independently by (531.1), that:

$$\exists w \forall q(w \models q \equiv q)$$

Suppose w_0 is an arbitrary such possible world, so that we know:

$$(\vartheta)\quad \forall q(w_0 \models q \equiv q)$$

Since w_0 is, by hypothesis, a possible world, it follows from our reductio assumption that $\neg w_0 \models p$. It also follows that w_0 is maximal (521). So by the definition of *Maximal*, $w_0 \models \neg p$. Hence, by (ϑ), it follows that $\neg p$, which contradicts our assumption that p. ⋈

(542.2) Our assumption is $\Diamond \exists w(w \models p)$ and we want to show $\exists w(w \models p)$. If we treat the restricted variable w in our assumption as singly restricted, our assumption becomes:

$$\Diamond\exists x(PossibleWorld(x)\,\&\,x \models p)$$

By BF$\Diamond$ (167.3), it follows from our assumption that:

$$\exists x\Diamond(PossibleWorld(x)\,\&\,x \models p)$$

Let b be an arbitrary such object, yielding $\Diamond(PossibleWorld(b)\,\&\,b \models p)$. Since the conjuncts of a possibly true conjunction are possible (162.3), it follows that:

(ξ) $\Diamond PossibleWorld(b)\,\&\,\Diamond b \models p$

Now by lemma (513.2), the first conjunct of (ξ) implies $PossibleWorld(b)$. Since $PossibleWorld(b)$, the facts (519) about the rigidity of truth at a possible world apply. So by (519.2), the second conjunct of (ξ) implies $b \models p$. Since we've established that $PossibleWorld(b)\,\&\,b \models p$, it follows that $\exists x(PossibleWorld(x)\,\&\,x \models p)$, i.e., $\exists w(w \models p)$. ⋈

(542.3) ($\rightarrow$) Assume p. By GEN, it suffices to show $\forall q(s \models q \equiv q) \rightarrow s \models p$. So assume $\forall q(s \models q \equiv q)$. As an instance of this latter assumption, we know $s \models p \equiv p$. Hence $s \models p$. ⋈

(542.4) Since (542.3) is modally strict, it follows by RN that:

$$\Box(p \rightarrow \forall s(\forall q(s \models q \equiv q) \rightarrow s \models p))$$

Hence by the K axiom (45.1), it follows that $\Box p \rightarrow \Box\forall s(\forall q(s \models q \equiv q) \rightarrow s \models p)$. ⋈

(542.5) To show $\Box\forall s\varphi \rightarrow \forall s\Box\varphi$, note that by our conventions for restricted variables, the formulas $\Box\forall s\varphi$ and $\forall s\Box\varphi$ are really shorthand for the formulas $\Box\forall s\varphi^s_x$ and $\forall s\Box\varphi^s_x$, respectively, for some free variable x in φ. So by eliminating the restricted variables, we have to show:

$$\Box\forall x(Situation(x) \rightarrow \varphi) \rightarrow \forall x(Situation(x) \rightarrow \Box\varphi)$$

So assume $\Box\forall x(Situation(x) \rightarrow \varphi)$. By the Converse Barcan Formula (167.2), it follows that:

(ϑ) $\forall x\Box(Situation(x) \rightarrow \varphi)$

Now to show $\forall x(Situation(x) \rightarrow \Box\varphi)$, it suffices by GEN to show $Situation(x) \rightarrow \Box\varphi$. So assume $Situation(x)$. By the rigidity of the notion of situation (469.1), it follows that:

(ξ) $\Box Situation(x)$

Moreover, it follows by applying $\forall$E to (ϑ) that:

(ζ) $\Box(Situation(x) \rightarrow \varphi)$

Hence, from (ξ) and (ζ) it follows that $\Box\varphi$, by the relevant instance of the K axiom (45.1). $\bowtie$

(542.6)[467] Assume $\forall w(w \models p)$. Now by (519.1), it is a modally strict theorem that $w \models p \equiv \Box w \models p$. So by a Rule of Substitution, $\forall w\Box(w \models p)$. But since we're working with bound restricted variables, the Barcan Formula holds. So $\Box\forall w(w \models p)$. $\bowtie$

(542.7) By the K axiom, it suffices to show $\Box(\forall w(w \models p) \to p)$. And by RN, it suffices to show $\forall w(w \models p) \to p$ by a modally strict proof. So assume $\forall w(w \models p)$. If we eliminate the restricted variable from our assumption, we know:

(ϑ) $\forall x(PossibleWorld(x) \to x \models p)$

Now by (531.2), we know *a fortiori* that $\exists x(PossibleWorld(x)\ \&\ Actual(x))$. Suppose a is an arbitrary such object, so that we know:

(ξ) $PossibleWorld(a)\ \&\ Actual(a)$

From (ϑ) and the first conjunct of (ξ), it follows that $a \models p$. From this and the second conjunct of (ξ), it follows by the definition of *Actual* that p. $\bowtie$

(543.1) ($\to$) By hypothetical syllogism from (542.1) and (542.2). ($\leftarrow$) Assume $\exists w(w \models p)$. Suppose w_1 is such a possible world, so that we know $w_1 \models p$. Then by the definition of possible world, $\Diamond\forall q(w_1 \models q \equiv q)$. By the Buridan$\Diamond$ formula (168.2), it follows that $\forall q\Diamond(w_1 \models q \equiv q)$. Hence, $\Diamond(w_1 \models p \equiv p)$ and, *a fortiori*, $\Diamond(w_1 \models p \to p)$. But we also know $w_1 \models p$ and so, by the rigidity of *truth at* (519.1), $\Box w_1 \models p$. Note that $\Diamond(\varphi \to \psi) \to (\Box\varphi \to \Diamond\psi)$ is the left-to-right direction of (162.4). So from $\Diamond(w_1 \models p \to p)$ and $\Box w_1 \models p$, it follows that $\Diamond p$. $\bowtie$

(543.2) The following proof uses (542.4) – (542.7). However, after the proof, we provide an alternative, simpler proof that appeals to (543.1) instead. ($\to$) Assume $\Box p$. It follows from this and (542.4) that:

$$\Box\forall s(\forall q(s \models q \equiv q) \to s \models p)$$

From this last fact and the version of the Converse Barcan Formula restricted to situations (542.5), it follows that:

(ξ) $\forall s\Box(\forall q(s \models q \equiv q) \to s \models p)$

Now we want to show: $\forall w(w \models p)$, i.e., $\forall s(PossibleWorld(s) \to s \models p)$ (treating w as doubly-restricted). By GEN, it suffices to show: $PossibleWorld(s) \to s \models p$. So assume $PossibleWorld(s)$. By the rigidity of *truth at* (473.2) for situations, it suffices to show $\Diamond s \models p$. By definition of *PossibleWorld* (512), we know:

[467] The following proof that takes advantage of w as a rigid restricted variable was contributed by Daniel West.

$\Diamond\forall q(s \models q \equiv q)$

But since (ξ) holds for every situation, it follows from (ξ) in particular that:

$\Box(\forall q(s \models q \equiv q) \rightarrow s \models p)$

But then, from our last two results it follows that $\Diamond s \models p$, by K$\Diamond$ (158.13).

($\leftarrow$) By hypothetical syllogism from (542.6) and (542.7). $\bowtie$

(543.2) (Alternative) By (529.1) and GEN, we know $\forall w((w \models \neg q) \equiv \neg(w \models q))$. But it is an simple exercise show that $\forall\alpha(\varphi \equiv \psi) \rightarrow (\exists\alpha\varphi \equiv \exists\alpha\psi)$. Hence it follows that:

(ϑ) $\exists w(w \models \neg q) \equiv \exists w \neg(w \models q)$

Then we may reason as follows:

1.	$\Diamond\neg q \equiv \exists w(w \models \neg q)$	Instance of (543.1)
2.	$\Diamond\neg q \equiv \exists w \neg(w \models q)$	From 1 and (ϑ)
3.	$\neg\Diamond\neg q \equiv \neg\exists w \neg(w \models q)$	From 2 and (88.4.b)
4.	$\Box q \equiv \neg\exists w \neg(w \models q)$	From 3 and Df$\Box$ (158.12)
5.	$\Box q \equiv \forall w(w \models q)$	From 4 and (103.3) $\bowtie$

(543.3) (Exercise)

(543.4) (Exercise)

(544.1) By (165.2), $\Box p \equiv \Diamond\Box p$. But by (543.1), $\Diamond\Box p \equiv \exists w(w \models \Box p)$. Hence $\Box p \equiv \exists w(w \models \Box p)$. $\bowtie$

(544.2) By (165.6), $\Box p \equiv \Box\Box p$. But by (543.2), $\Box\Box p \equiv \forall w(w \models \Box p)$. Hence $\Box p \equiv \forall w(w \models \Box p)$. $\bowtie$

(544.3) – **(544.4)** (Exercises)

(545.1) ($\rightarrow$) Assume $w \models (p \,\&\, q)$. Since w is by (528) n-modally closed for $n \geq 1$, we know that the following unary instances of the definition (527) of n-modal closure obtain:

$\forall r_1 \forall r_2((w \models r_1 \,\&\, (r_1 \Rightarrow r_2)) \rightarrow w \models r_2)$

Substituting $p \,\&\, q$ for r_1 and p for r_2, we obtain:

$(w \models (p \,\&\, q) \,\&\, ((p \,\&\, q) \Rightarrow p)) \rightarrow w \models p$

And substituting $p \,\&\, q$ for r_1 and q for r_2, we obtain:

$(w \models (p \,\&\, q) \,\&\, ((p \,\&\, q) \Rightarrow q)) \rightarrow w \models q$

Since the first conjunct of each antecedent is just our assumption and the second conjunct of each antecedent follows by an easily proved fact (exercise), we can derive, respectively, $w \models p$ and $w \models q$. Hence $(w \models p) \,\&\, (w \models q)$.

($\leftarrow$) Assume $(w \models p) \,\&\, (w \models q)$. Since w is by (528) n-modally closed for $n \geq 1$, we know that the following binary instance of the definition (527) of n-modal closure obtains:

$$\forall r_1 \forall r_2 \forall r_3 ((w \models r_1 \;\&\; w \models r_2 \;\&\; ((r_1 \,\&\, r_2) \Rightarrow r_3)) \rightarrow w \models r_3)$$

Substituting p for r_1, q for r_2, and $p \,\&\, q$ for r_3, we obtain:

$$(w \models p \;\&\; w \models q \;\&\; ((p \,\&\, q) \Rightarrow (p \,\&\, q))) \rightarrow w \models (p \,\&\, q)$$

Now the first two conjuncts of the antecedent follow from our initial assumption, and the third conjunct of the antecedent is easily established (exercise). It follows that $w \models (p \,\&\, q)$. ⋈

(545.2) ($\rightarrow$) Assume both $w \models (p \rightarrow q)$ and $w \models p$. Clearly, it is a modally strict theorem that $((p \rightarrow q) \,\&\, p) \rightarrow q$. Hence by RN and the definition of necessary implication, $((p \rightarrow q) \,\&\, p) \Rightarrow q$. So by the relevant instance of the fact that possible worlds are modally closed, it follows that $w \models q$.

($\leftarrow$) Assume $(w \models p) \rightarrow (w \models q)$. Assume, for reductio, that $\neg w \models (p \rightarrow q)$. Then by (529.1), $w \models \neg(p \rightarrow q)$. But since it is a modally strict theorem that $\neg(p \rightarrow q) \rightarrow (p \,\&\, \neg q)$ (88.1.b), it follows by RN and the definition of $\Rightarrow$ (524.1) that $\neg(p \rightarrow q) \Rightarrow (p \,\&\, \neg q)$. Hence $w \models (p \,\&\, \neg q)$, by the fact that possible worlds are 1-modally closed (528). So by (545.1):

(ϑ) $(w \models p) \,\&\, (w \models \neg q)$

The first conjunct of (ϑ) and our initial assumption jointly imply $w \models q$. But the second conjunct of (ϑ) implies $\neg w \models q$, by (529.1). Contradiction. ⋈

(545.3) In what follows, we use $w \not\models p$ to abbreviate $\neg w \models p$.

($\rightarrow$) For conditional proof, assume:

(a) $w \models (p \vee q)$

Suppose, for reductio, that $\neg(w \models p \vee w \models q)$, i.e.,

(b) $(w \not\models p) \,\&\, (w \not\models q)$

Then, by (529.1), it follows from the first conjunct of (b) that:

(c) $w \models \neg p$

Independently, starting with disjunctive syllogism (86.4.b), we may derive $((p \vee q) \,\&\, \neg p) \rightarrow q$ by a modally strict proof. Hence, by RN:

(d) $\Box(((p \vee q) \,\&\, \neg p) \to q)$

By definition of $\Rightarrow$ (524), (d) implies:

(e) $((p \vee q) \,\&\, \neg p) \Rightarrow q$

But by the fact that worlds are 2-modally closed (528), it follows from (a), (c), and (e) that $w \models q$. But this contradicts the second conjunct of (b).

($\leftarrow$) (Exercise) ⋈

(545.4) (Exercise)

(545.5) If we eliminate the restricted variable, then the theorem to be proved is: $PossibleWorld(s) \to ((s \models \forall\alpha\varphi) \equiv \forall\alpha(s \models \varphi))$. Our proof strategy is:[468]

(A) Establish $\Box(\forall p(s \models p \equiv p) \to ((s \models \forall\alpha\varphi) \equiv \forall\alpha(s \models \varphi)))$, by applying, to (511.3), our expanded RN (341.3.a) and the fact that the free restricted variable s is rigid.

(B) Infer from (A) that $\Diamond\forall p(s \models p \equiv p) \to \Diamond((s \models \forall\alpha\varphi) \equiv \forall\alpha(s \models \varphi))$, by the relevant instance of K$\Diamond$ (158.13).

(C) Conclude from (B) that $PossibleWorld(s) \to \Diamond((s \models \forall\alpha\varphi) \equiv \forall\alpha(s \models \varphi))$ by definition (512).

(D) Independently show that $\Diamond((s \models \forall\alpha\varphi) \equiv \forall\alpha(s \models \varphi)) \to ((s \models \forall\alpha\varphi) \equiv \forall\alpha(s \models \varphi))$.

(E) Our theorem then follows by hypothetical syllogism from (C) and (D).

It remains only to show (D). But to show (D), it suffices, by metarule (166.2) when Γ is empty, to show that the following is a modally strict theorem:

(F) $((s \models \forall\alpha\varphi) \equiv \forall\alpha(s \models \varphi)) \to \Box((s \models \forall\alpha\varphi) \equiv \forall\alpha(s \models \varphi))$

Note that we can prove (F) by modally strict means if we establish that the two formulas $s \models \forall\alpha\varphi$ and $\forall\alpha(s \models \varphi)$ both exhibit modal collapse, i.e., that:

(ϑ) $\Box((s \models \forall\alpha\varphi) \to \Box s \models \forall\alpha\varphi)$

(ξ) $\Box(\forall\alpha(s \models \varphi) \to \Box\forall\alpha(s \models \varphi))$

For then, the relevant instance of (172.5) and the T schema (45.2) jointly imply that (F) is a theorem. So it remains to show (ϑ) and (ξ).

(ϑ) Assume $s \models \forall\alpha\varphi$. Then, by (473.1), $\Box s \models \forall\alpha\varphi$. So, by conditional proof, $(s \models \forall\alpha\varphi) \to \Box s \models \forall\alpha\varphi$ is a modally strict theorem. Since the free restricted

[468] I'm indebted to Daniel Kirchner, who developed the following proof sketch in one of our working sessions on object theory.

variables are rigid, we can conclude $\Box((s \models \forall\alpha\varphi) \rightarrow \Box s \models \forall\alpha\varphi)$ by applying expanded RN (341.3.a).

(ξ) Assume $\forall\alpha(s \models \varphi)$. Then by $\forall$E, $s \models \varphi$. So again by (473.1), $\Box s \models \varphi$. Since α isn't free in our assumption, it follows by GEN that $\forall\alpha\Box s \models \varphi$. Hence by BF (167.1), $\Box\forall\alpha(s \models \varphi)$. So, by conditional proof, it is a modally strict theorem that $\forall\alpha(s \models \varphi) \rightarrow \Box\forall\alpha(s \models \varphi)$. Since the free restricted variables are rigid, we can conclude $\Box(\forall\alpha(s \models \varphi) \rightarrow \Box\forall\alpha(s \models \varphi))$ by applying expanded RN (341.3.a). ⋈

(545.6) (Exercise)

(545.7) Assume $w \models \Box p$. Since $\Box p \rightarrow p$ is an instance of the T schema axiom, we may conclude by RN that $\Box(\Box p \rightarrow p)$. Since possible worlds are modally closed, it follows that $w \models p$. So by (519.1), $\Box w \models p$. ⋈

(545.8) See footnote 288.

(545.9) Assume $\Diamond w \models p$. Assume, for reductio, that $\neg w \models \Diamond p$. By the coherency of possible worlds (529.1), it follows that $w \models \neg\Diamond p$. But since worlds are modally closed and $\neg\Diamond p$ is necessarily equivalent to $\Box\neg p$, it follows that $w \models \Box\neg p$. Hence, by (545.7), $\Box w \models \neg p$. Since the coherency of possible worlds is a modally strict theorem, it follows by a Rule of Substitution that $\Box\neg w \models p$. But this is equivalent to $\neg\Diamond w \models p$. Contradiction. ⋈

(545.10) Since there are contingently false propositions (217.2), let r be such a proposition, so that we know by (213.1) that $\neg r \,\&\, \Diamond r$. By the T$\Diamond$ schema, it follows that $\Diamond(\neg r \,\&\, \Diamond r)$. So by a fundamental theorem of possible world theory (543.1), $\exists w(w \models (\neg r \,\&\, \Diamond r))$. Let w_1 be such a possible world, so that we know $w_1 \models (\neg r \,\&\, \Diamond r)$. Then by (545.1), it follows that:

(ϑ) $w_1 \models \neg r$

(ξ) $w_1 \models \Diamond r$

From (ϑ) and the coherence of possible worlds, it follows that $\neg w_1 \models r$. Hence by (519.5), $\Box\neg w_1 \models r$, which implies $\neg\Diamond w_1 \models r$. Conjoining ($\xi$) with this last result yields $(w_1 \models \Diamond r) \,\&\, \neg\Diamond w_1 \models r$. Hence, $\exists w\exists p((w \models \Diamond p) \,\&\, \neg\Diamond w \models p)$. ⋈

(547.1) Assume $\exists p\mathit{ContingentlyTrue}(p)$. Let p_1 be an arbitrary such proposition, so that, by definition, we know both p_1 and $\Diamond\neg p_1$. Now if we instantiate a Fundamental Theorem of Possible World Theory (543.1) to $\neg p_1$, we know $\Diamond\neg p_1 \equiv \exists w(w \models \neg p_1)$. So from this and the second conjunct of our assumption, it follows that $\exists w(w \models \neg p_1)$. Let w_1 be such a possible world, so that we know $w_1 \models \neg p_1$. Now suppose, for reductio, that $\mathit{Actual}(w_1)$. Then by definition of *Actual*, every proposition true at w_1 is true. Hence $\neg p_1$, which is a contradiction. So $\neg\mathit{Actual}(w_1)$ and, hence, $\exists w\neg\mathit{Actual}(w)$. ⋈

(547.2) (Exercise)

(547.3) (The proof is given in the text.)

(547.4) Since $\exists! w Actual(w)$ (531.2), suppose w_1 is such a possible world, so that we know $Actual(w_1)$, among other things. Independently, by (547.3), we know $\exists w \neg Actual(w)$. Let w_2 be such a possible world, so that we know $\neg Actual(w_2)$. Since w_1 is actual and w_2 is not, it follows that $w_1 \neq w_2$. Hence $\exists w \exists w'(w \neq w')$. ⋈

(549)★ Our strategy is to prove a stronger claim, $\Box\exists p(p \,\&\, \forall q(q \to \Box(p \to q)))$, so that our resulte follows by the T schema.[469] Then by a fundamental theorem of world theory (543.2), it suffices to show:

$$\forall w\big(w \models \exists p(p \,\&\, \forall q(q \to \Box(p \to q)))\big)$$

Since GEN is valid for restricted variables, it suffices to show:

$$w \models \exists p(p \,\&\, \forall q(q \to \Box(p \to q)))$$

By (545.6), this is equivalent to:

$$\exists p\big(w \models (p \,\&\, \forall q(q \to \Box(p \to q)))\big)$$

Now w is a fixed, but arbitrary possible world, so let p_1 be the proposition $\forall p(p \equiv w \models p)$. To establish that p_1 is a witness to the above, we have to show:

[469] In West's original communication (01 January 2023), he presented a non-modally strict proof, which goes as follows. Let p_1 be the formula $\forall p(p \equiv w_\alpha \models p)$, where w_α is the actual world. Note that p_1 is a theorem and hence true (536.2)★. So it remains to show that p_1 necessarily implies every true proposition, i.e.,

$$\forall q(q \to \Box(p_1 \to q))$$

By GEN, it suffices to show, $q \to \Box(p_1 \to q)$. So assume q, to show $\Box(p_1 \to q)$. By a fundamental theorem of possible world theory (543.2), it suffices to show:

$$\forall w(w \models (p \to q))$$

Again, by GEN (which is valid for restricted variables), it suffices to show $w \models (p_1 \to q)$. By the equivalence (545.2), it suffices to show $(w \models p_1) \to (w \models q)$. So assume $w \models p_1$, to show $w \models q$. Then by definition of p_1:

$$w \models \forall p(p \equiv w_\alpha \models p)$$

Applying (545.5), this assumption is equivalent to:

$$\forall p(w \models (p \equiv w_\alpha \models p))$$

Instantiating to q, it follows that:

$$w \models (q \equiv w_\alpha \models q)$$

Thus by (545.4):

$$(w \models q) \equiv w \models (w_\alpha \models q)$$

Since we're trying to show the left side, it remains only to prove $w \models (w_\alpha \models q)$. But q is true, by hypothesis, and so we know $w_\alpha \models q$ as an instance of (536.2)★. Hence, it follows that $\Box w_\alpha \models q$, by the rigidity of truth at a world (519.1). Then by a fundamental theorem (543.2), $\forall w(w \models (w_\alpha \models q))$. Instantiating to w, it follows that $w \models (w_\alpha \models q)$. ⋈

$$w \models (p_1 \,\&\, \forall q(q \to \Box(p_1 \to q)))$$

By (545.1), it suffices to show:

$$(w \models p_1) \,\&\, w \models \forall q(q \to \Box(p_1 \to q))$$

We prove the conjuncts in turn.

(A) Show $w \models p_1$. By definition of p_1, we have to show $w \models \forall p(p \equiv w \models p)$. By (545.5), it suffices to show $\forall p(w \models (p \equiv w \models p))$. So by GEN, we need only show $w \models (p \equiv w \models p)$, i.e., by (545.4), $(w \models p) \equiv w \models (w \models p)$. But this is an almost immediate consequence of the following, more general, modally strict lemma:

Lemma. $\forall p \forall w \forall w'((w \models p) \equiv w' \models (w \models p))$

Proof. By GEN, it suffices to show $(w \models p) \equiv w' \models (w \models p)$. ($\to$) Assume $w \models p$. Then by the rigidity of truth at a world (519.1), $\Box w \models p$. By a fundamental theorem of world theory, $\forall w'(w' \models (w \models p))$. Instantiating to w', we obtain $w' \models (w \models p)$. ($\leftarrow$) Assume $w' \models (w \models p)$. Clearly then $\exists w'(w' \models (w \models p))$. Hence, by a fundamental theorem of world theory, $\Diamond w \models p$. So by (519.2), $w \models p$.

(B) Show $w \models \forall q(q \to \Box(p_1 \to q))$. By (545.5), it suffices to show:

$$\forall q\big(w \models (q \to \Box(p_1 \to q))\big)$$

and by GEN:

$$w \models (q \to \Box(p_1 \to q))$$

But is last is equivalent, by (545.2), to:

$$(w \models q) \to w \models \Box(p_1 \to q)$$

So assume $w \models q$. Though we want to show $w \models \Box(p_1 \to q)$, note that if we show $\Box(p_1 \to q)$, then by (544.2) we could infer $\forall w'(w' \models \Box(p_1 \to q))$ and by instantiating to w, we would then have $w \models \Box(p_1 \to q)$. So, to show $\Box(p_1 \to q)$, suppose not, for reductio. Then $\Diamond\neg(p_1 \to q)$ and by classical modal reasoning, $\Diamond(p_1 \,\&\, \neg q)$. So by a fundamental theorem of possible world theory (543.1), $\exists w'(w' \models (p_1 \& \neg q))$. Suppose w_1 is such a world, so that we know $w_1 \models (p_1 \& \neg q)$. Then by (545.1), we know both $w_1 \models p_1$ and $w_1 \models \neg q$. The first of these, by definition of p_1, comes to: $w_1 \models \forall p(p \equiv w \models p)$. So by (545.5):

$$\forall p(w_1 \models (p \equiv w \models p))$$

By (545.4) and a Rule of Substitution, we may infer:

$$(\vartheta) \quad \forall p\big((w_1 \models p) \equiv w_1 \models (w \models p)\big)$$

Now recall the modally strict Lemma proved above. Commute the quantifiers in that Lemma so that we have:

$$\forall w \forall w' \forall p((w \models p) \equiv w' \models (w \models p))$$

Instantiate the first quantifier to w and the second to w_1 and we therefore know:

$$(\xi)\ \ \forall p\big((w \models p) \equiv w_1 \models (w \models p)\big)$$

Note that (ϑ) has the form $\forall p(\varphi \equiv \psi)$ and that (ξ) has the form $\forall p(\chi \equiv \psi)$. So we may infer something of the form $\forall p(\varphi \equiv \chi)$, i.e., $\forall p((w_1 \models p) \equiv w \models p)$. Hence, $w_1 = w$, by (516). But then from our earlier result that $w_1 \models \neg q$ it follows that $w \models \neg q$. So by the coherency of possible worlds (529.1), $\neg w \models q$, which contradicts the assumption that $w \models q$. ⋈[470]

(551) From the fact that $[\lambda y p]{\downarrow}$ (39.2) and β-Conversion, we know that $[\lambda y p]x \equiv p$ is a modally strict theorem. Hence, by RN, $\Box([\lambda y\, p]x \equiv p)$, and by (158.4), $\Box([\lambda y\, p]x \to p)$ and $\Box(p \to [\lambda y\, p]x)$. Hence by definition of $\Rightarrow$ (524), we know both:

$$(\vartheta)\ \ [\lambda y\, p]x \Rightarrow p$$

$$(\zeta)\ \ p \Rightarrow [\lambda y\, p]x$$

We now establish our theorem by arguing for both directions. ($\rightarrow$) Assume $w \models p$. Then by (ζ), the fact that possible worlds are 1-modally closed (528), and the definition of n-modal closure (527), it follows that $w \models [\lambda y\, p]x$. ($\leftarrow$) By analogous reasoning, from (ϑ). ⋈

(552.1) Assume, for reductio, $\exists w \exists p(w \models (p \& \neg p))$. Let w_1 and p_1 be an arbitrary such world and proposition, respectively, so that we know $w_1 \models (p_1 \,\&\, \neg p_1)$. Hence, $\exists w(w \models (p_1 \,\&\, \neg p_1))$. Then by fundamental theorem (543.1), it follows that $\Diamond(p_1 \,\&\, \neg p_1)$, contradicting the fact, provable from (84), RN, and (162.1), that $\neg\Diamond(p_1 \,\&\, \neg p_1)$. ⋈

(552.2) By (552.1), we know that $\neg\exists w \exists p(w \models (p \,\&\, \neg p))$. This is equivalent to $\forall w \neg\exists p(w \models (p \,\&\, \neg p))$, and so by $\forall$E, $\neg\exists p(w \models (p \,\&\, \neg p))$. Hence, by (77.3), $\exists p(w \models (p \,\&\, \neg p)) \to \forall q(w \models q)$. ⋈

(553) (Exercise)

(556.1) – **(556.2)** (Exercises)

(558.1) – **(558.2)** (Exercises)

---

[470] This preserves, almost intact, the proof described in Daniel West's personal communication of 02 January 2023.

(558.3) By definition of $\circ_w p$ (558.1), theorem (261.2), and the fact that the formula $w \models (q \equiv p)$ is a rigid condition on properties (558.2). ⋈

(558.4) – **(558.5)** (Exercises)

(558.6) By definition (558.3), we know the following respective facts about $\circ_w p$ and $\circ_w q$:

(ϑ) $\forall r(\circ_w p \models r \equiv w \models (r \equiv p))$

(ξ) $\forall r(\circ_w q \models r \equiv w \models (r \equiv q))$

($\rightarrow$) Assume $\circ_w p = \circ_w q$. Now (558.5) tells us that $\circ_w p \models p$. Hence $\circ_w q \models p$. This implies $w \models (p \equiv q)$ by (ξ).

($\leftarrow$) Assume $w \models (p \equiv q)$. Since $\circ_w p$ and $\circ_w q$ are both situations, it suffices to show $\forall r((\circ_w p \models r) \equiv (\circ_w q \models r))$, by (474). So, by GEN, we show $(\circ_w p \models r) \equiv (\circ_w q \models r)$. ($\rightarrow$) Assume $\circ_w p \models r$. Then by (ϑ), we know: $w \models (r \equiv p)$. Since we also know $w \models (p \equiv q)$ by assumption, it follows that $w \models (r \equiv q)$, from the facts that (a) $\Box(((r \equiv p) \& (p \equiv q)) \rightarrow (r \equiv q))$ and (b) w is 2-modally closed (528). Hence, $\circ_w q \models r$, by (ξ). ($\leftarrow$) By analogous reasoning. ⋈

(560.1) – **(560.5)** (Exercises)

(561.1) By (558.3), we know:

(ϑ) $\forall q(\circ_w p \models q \equiv w \models (q \equiv p))$

Moreover, theorem (560.3) is:

(ξ) $\forall q(\top_w \models q \equiv w \models q)$

We now prove both directions of our biconditional theorem.

($\rightarrow$) Assume $w \models p$. By (474), it suffices to show $\forall r(\circ_w p \models r \equiv \top_w \models r)$. By GEN, it suffices to show $\circ_w p \models r \equiv \top_w \models r$.

> ($\rightarrow$) Assume $\circ_w p \models r$. Then by (ϑ), it follows that $w \models (r \equiv p)$. But from this last fact, our assumption that $w \models p$, and the easily established fact that $((r \equiv p) \& p) \Rightarrow r$, it follows that $w \models r$, by the 2-modal closure of possible worlds (528). Hence, by (560.3), $\top_w \models r$.
>
> ($\leftarrow$) Assume $\top_w \models r$. Then by (560.3), $w \models r$. But from this last fact, our assumption that $w \models p$, and the easily established fact that $(r \& p) \Rightarrow (r \equiv p)$, it follows that $w \models (r \equiv p)$, again by the 2-modal closure of possible worlds. It follows by (ϑ) that $\circ_w p \models r$.

($\leftarrow$) Assume $\circ_w p = \top_w$. Independently, by (558.5), we know $\circ_w p \models p$. Hence, $\top_w \models p$, by the substitution of identicals. But then by (560.3), $w \models p$. ⋈

(561.2) (Exercise)

(562.1) By applying GEN to theorem (561.1), we know:

$$\forall w(w \models p \equiv \circ_w p = \top_w)$$

It follows from this, by a law of quantification theory (99.3), that:

(ϑ) $\forall w(w \models p) \equiv \forall w(\circ_w p = \top_w)$

But it is a Fundamental Theorem of Possible World Theory (543.2) that:

$$\Box p \equiv \forall w(w \models p)$$

From this and (ϑ), it follows that $\Box p \equiv \forall w(\circ_w p = \top_w)$. ⋈

(562.2) – **(562.4)** (Exercises)

(564.1) – **(564.2)** (Exercises)

(566.1) (Exercise)

(566.2) By GEN, it suffices to show $\varphi \rightarrow \Box\varphi$. So assume φ, that is, $w \models \forall y(Fy \equiv Gy)$. Then by the rigidity of truth at a world (519.1), it follows that $\Box w \models \forall y(Fy \equiv Gy)$, i.e., $\Box\varphi$.

(566.3) By definition of $\epsilon_w G$, (563), theorem (261.2) and the fact (566.2) that $w \models \forall y(Fy \equiv Gy)$ is a rigid condition on properties. ⋈

(566.4) – **(566.5)** (Exercises)

(567.1) Assume $ExtensionAtOf(x, w, G)$ and $ExtensionAtOf(y, w, H)$. By the definition of $ExtensionAtOf$ (563), we therefore know:

(a) $A!x \,\&\, G{\downarrow} \,\&\, \forall F(xF \equiv w \models \forall z(Fz \equiv Gz))$

(b) $A!y \,\&\, H{\downarrow} \,\&\, \forall F(yF \equiv w \models \forall z(Fz \equiv Hz))$

($\rightarrow$) Assume $x = y$. Then by Rule =E, it follows from (a) that:

(c) $A!y \,\&\, G{\downarrow} \,\&\, \forall F(yF \equiv w \models \forall z(Fz \equiv Gz))$

Hence, by (99.11) and (99.10), the third conjuncts of (b) and (c) imply:

(d) $\forall F[w \models \forall z(Fz \equiv Hz) \equiv w \models \forall z(Fz \equiv Gz)]$

Now if we instantiate (d) to G, we know:

(e) $w \models \forall z(Gz \equiv Hz) \equiv w \models \forall z(Gz \equiv Gz)$

But the right side of (e) is easily derived: from the tautology $Gz \equiv Gz$, we obtain $\forall z(Gz \equiv Gz)$, by GEN. Since this is a modally strict theorem, we obtain $\Box\forall z(Gz \equiv Gz)$. So by a Fundamental Theorem of Possible World Theory (543.2), it follows that $\forall w'(w' \models \forall z(Gz \equiv Gz))$. Instantiating to w, we obtain the right side of (e). So the left side of (e), $w \models \forall z(Gz \equiv Hz)$, follows by biconditional syllogism.

($\leftarrow$) Assume:

(f) $w \models \forall z(Gz \equiv Hz)$

Since we know $A!x$ and $A!y$ by the left conjuncts of (a) and (b), it suffices by theorem (245.2) to show $\forall F(xF \equiv yF)$, and by GEN, that $xF \equiv yF$:

($\rightarrow$) Assume xF. Then by the right conjunct of (a), it follows that

(g) $w \models \forall z(Fz \equiv Gz)$

Now, independently, if we apply RN to the modally strict theorem (99.10) and apply the definition $\Rightarrow$ (524), then we know:

(h) $(\forall z(Fz \equiv Gz) \,\&\, \forall z(Gz \equiv Hz)) \Rightarrow \forall z(Fz \equiv Hz)$

Hence, from (g), (f), and (h), it follows by the 2-modal closure of possible worlds (528) that $w \models \forall z(Fz \equiv Hz)$. Hence, by the right conjunct of (b), yF.

($\leftarrow$) (Exercise) ⋈

(567.2) Since $\epsilon_w F\downarrow$ and $\epsilon_w G\downarrow$, we can instantiate $\epsilon_w F$ and $\epsilon_w G$ into world-relativized pre-Law V (567.1). Simultaneously substituting $\epsilon_w F$ for x, F for G, $\epsilon_w G$ for y, and G for H, we obtain:

$$(\mathit{ExtensionAtOf}(\epsilon_w F, w, F) \,\&\, \mathit{ExtensionAtOf}(\epsilon_w G, w, G)) \rightarrow \\ (\epsilon_w F = \epsilon_w G \;\equiv\; w \models \forall z(Fz \equiv Gz))$$

But we also know both conjuncts of the antecedent, by (566.4). Hence, $\epsilon_w F = \epsilon_w G \equiv w \models \forall z(Fz \equiv Gz)$. ⋈

(569.1) Without loss of generality, suppose that y doesn't occur free in φ; if φ is a formula in which y occurs free, then we can still establish our theorem by appealing to alphabetic variants of the principles used below. Our strategy is to prove the theorem by conditional proof, i.e., assume $[\lambda x\, \varphi]\downarrow$ and show $[\lambda x\, w \models \varphi]\downarrow$, where w is some fixed but arbitrary possible world. Before we begin the proof, we first establish a Lemma about our fixed but arbitrary world w:

Lemma: $\Box[\lambda x\, \varphi]\downarrow \rightarrow \Box\forall x \forall y(\forall F(Fx \equiv Fy) \rightarrow (w \models \varphi \equiv w \models \varphi^y_x))$

Proof. By Rule RM, it suffices to show the following by a modally strict proof:

$$[\lambda x\, \varphi]\downarrow \rightarrow \forall x \forall y(\forall F(Fx \equiv Fy) \rightarrow (w \models \varphi \equiv w \models \varphi^y_x))$$

So assume $[\lambda x\, \varphi]\downarrow$. Then by the Corollary to Kirchner's Theorem (272), it follows that:

$$\forall x \forall y(\forall F(Fx \equiv Fy) \rightarrow \Box(\varphi \equiv \varphi^y_x))$$

Hence, by ∀E:

(ϑ) $\forall F(Fx \equiv Fy) \rightarrow \Box(\varphi \equiv \varphi^y_x)$

But it is a fact about the consequent that:

(ξ) $\Box(\varphi \equiv \varphi^y_x) \rightarrow (w \models \varphi \equiv w \models \varphi^y_x)$

This is established by the following hypothetical syllogism chain:

$$\begin{array}{lcll} \Box(\varphi \equiv \varphi^y_x) & \rightarrow & \forall w'(w' \models (\varphi \equiv \varphi^y_x)) & \text{by (543.2)} \\ & \rightarrow & w \models (\varphi \equiv \varphi^y_x) & \text{by } \forall\text{E} \\ & \rightarrow & w \models \varphi \equiv w \models \varphi^y_x & \text{by (545.4)} \end{array}$$

So from (ϑ) and (ξ) if follows that:

$$\forall F(Fx \equiv Fy) \rightarrow (w \models \varphi \equiv w \models \varphi^y_x)$$

Since x and y are not free in our assumption, it follows by GEN that:

$$\forall x \forall y(\forall F(Fx \equiv Fy) \rightarrow (w \models \varphi \equiv w \models \varphi^y_x))$$

Now using this Lemma, we prove our theorem. Assume $[\lambda x \varphi]\downarrow$. Then by (106), $\Box[\lambda x\, \varphi]\downarrow$. So by our Lemma:

$$\Box\forall x \forall y(\forall F(Fx \equiv Fy) \rightarrow (w \models \varphi \equiv w \models \varphi^y_x))$$

But $w \models \varphi^y_x$ is the same formula as $(w \models \varphi)^y_x$. Hence:

$$\Box\forall x \forall y(\forall F(Fx \equiv Fy) \rightarrow (w \models \varphi \equiv (w \models \varphi)^y_x))$$

So, by an instance of the right-to-left direction of Kirchner's Theorem (271.1), in which we set φ in Kirchner's Theorem to $w \models \varphi$, it follows that $[\lambda x\, w \models \varphi]\downarrow$. ⋈

(569.2) The proof is by cases:

$n = 0$. $[\lambda\ w \models \varphi]$ is a core λ-expression and so by (39.2), $[\lambda\ w \models \varphi]\downarrow$. Hence, $[\lambda\, \varphi]\downarrow \rightarrow [\lambda\ w \models \varphi]\downarrow$, by the truth of the consequent.

$n = 1$. By (569.1).

$n \geq 2$. (Exercise)

(569.3) See the discussion in the text as to why this is an instance of axiom (39.2). ⋈

(573.1) We prove this for the cases $n \geq 1$ and $n = 0$.

Case $n \geq 1$. By (569.3), $[\lambda x_1 \ldots x_n w \models F^n x_1 \ldots x_n]\downarrow$. So we may reason with this relation as follows:

$$\begin{array}{rcll} F^n_w x_1 \ldots x_n & \equiv & [\lambda x_1 \ldots x_n \, w \models F^n x_1 \ldots x_n] x_1 \ldots x_n & \text{by (570.1), Rules } =_{df}\text{E}, =_{df}\text{I} \\ & \equiv & w \models F^n x_1 \ldots x_n & \text{by } \beta\text{-Conversion} \end{array}$$

Case $n = 0$. (Exercise) ⋈

(570.2) By (569.3), we know $[\lambda x_1 \ldots x_n \, w \models F x_1 \ldots x_n]\downarrow$ $(n \geq 0)$. So by our theory of definitions and the definition of F^n_w (570.1), $F^n_w\downarrow$. So by two applications of GEN, $\forall F^n \forall w (F^n_w \downarrow)$. ⋈

(573.2) We prove this for the cases $n \geq 1$ and $n = 0$.

Case $n \geq 1$. By the definition of *rigid* (571.1), we have to show:

(A) $G^n_w\downarrow$

(B) $\Box \forall x_1 \ldots \forall x_n (G^n_w x_1 \ldots x_n \rightarrow \Box G^n_w x_1 \ldots x_n)$

(A) By theorem (570.2).

(B) By RN and GEN, we have to show: $G^n_w x_1 \ldots x_n \rightarrow \Box G^n_w x_1 \ldots x_n$. So assume $G^n_w x_1 \ldots x_n$. Then by (573.1), $w \models G^n x_1 \ldots x_n$. But possible worlds necessarily encode any propositions true at them; i.e., $w \models p \rightarrow \Box w \models p$ is the left-to-right direction of theorem (519.1). Hence $\Box w \models G^n x_1 \ldots x_n$. But by a Rule of Substitution and the fact that (573.1) is a modally strict theorem, it follows that $\Box G^n_w x_1 \ldots x_n$.

Case $n = 0$. (Exercise) ⋈

(573.3) Note that we also include an alternative proof after the first proof. The alternative (somewhat simpler) proof was contributed by Daniel Kirchner. But both are of interest, since they approach the problem in two completely different ways.

We prove our theorem by cases $n \geq 1$ and $n = 0$.

Case $n \geq 1$. By the T schema, it suffices to prove the stronger claim:

$$\Box \exists F^n (\mathit{Rigidifies}(F^n, G^n))$$

By (543.2), it suffices to show:

$$\forall w (w \models \exists F^n (\mathit{Rigidifies}(F^n, G^n)))$$

By GEN, it suffices to show:

$$w \models \exists F^n (\mathit{Rigidifies}(F^n, G^n))$$

By (545.6), it suffices to show:

$$\exists F^n (w \models \mathit{Rigidifies}(F^n, G^n)$$

Let's then show that G^n_w is a witness to this existential claim, i.e., that:

(ϑ) $w \models Rigidifies(G^n_w, G^n)$

Note that the definition of *Rigidifies* (571.2) yields the necessary equivalence:

$$\Box(Rigidifies(G^n_w, G^n) \equiv (Rigid(G^n_w) \,\&\, \forall x_1 \ldots \forall x_n(G^n_w x_1 \ldots x_n \equiv G^n x_1 \ldots x_n)))$$

So since worlds are 2-modally closed (528), we can establish (ϑ) if we can show:[471]

$$w \models (Rigid(G^n_w) \,\&\, \forall x_1 \ldots \forall x_n(G^n_w x_1 \ldots x_n \equiv G^n x_1 \ldots x_n))$$

So by (545.1), it suffices to show:

(A) $w \models Rigid(G^n_w)$

(B) $w \models \forall x_1 \ldots \forall x_n(G^n_w x_1 \ldots x_n \equiv G^n x_1 \ldots x_n)$

(A) Theorem (573.2) is modally strict. Since the free restricted variable is rigid, it follows by expanded RN (341.3.a) that $\Box Rigid(G^n_w)$. A fundamental theorem of world theory (543.2) then implies $\forall w'(w' \models Rigid(G^n_w))$. Instantiating to w, we have $w \models Rigid(G^n_w)$.

(B) By (545.5), it suffices to show:

$$\forall x_1 \ldots \forall x_n(w \models (G^n_w x_1 \ldots x_n \equiv G^n x_1 \ldots x_n))$$

So by GEN, we need only show:

$$w \models (G^n_w x_1 \ldots x_n \equiv G^n x_1 \ldots x_n)$$

And by (545.4), it suffices to show:

(ζ) $w \models G^n_w x_1 \ldots x_n \equiv w \models G^n x_1 \ldots x_n$

To prove this, we use the following modally strict theorem and its Corollary:

> *Lemma*: $G^n_w x_1 \ldots x_n \equiv \Box G^n_w x_1 \ldots x_n$
> *Proof*. (→) If $G^n_w x_1 \ldots x_n$, then since G^n_w is rigid (573.2), it follows by the definition of rigidity (571.1) that $\Box G^n_w x_1 \ldots x_n$. (←) This is an instance of the T schema.
>
> *Corollary*: $\Diamond G^n_w x_1 \ldots x_n \equiv G^n_w x_1 \ldots x_n$
> *Proof*. (→) Since $G^n_w x_1 \ldots x_n \to \Box G^n_w x_1 \ldots x_n$ is a modally strict theorem implied by our Lemma, it follows by Rule (166.2) that $\Diamond G^n_w x_1 \ldots x_n \equiv G^n_w x_1 \ldots x_n$. (←) This is an instance of the T◊ schema.

[471] We have to appeal to the modal closure of possible worlds here, since one can't validly substitute necessarily equivalent properties for one another in this context. Given that possible worlds are situations and, hence, abstract, (ϑ) becomes, by definition, necessarily equivalent to an encoding claim of the form $x[\lambda y\, \varphi]$, where y isn't free in φ. If ψ is necessarily equivalent to φ, we can't validly substitute ψ for φ in $x[\lambda y\, \varphi]$ to conclude $x[\lambda y\, \psi]$. It's crucial here that we're dealing with possible worlds. For then we know that $w \models \varphi$ implies $w \models \psi$ when φ and ψ are necessarily equivalent, by applying (528).

With this Lemma, we can prove both directions of (ζ).

($\rightarrow$) Assume $w \models G^n_w x_1 \dots x_n$. Hence $\exists w(w \models G^n_w x_1 \dots x_n)$, and so by a fundamental theorem of world theory (543.1), $\Diamond G^n_w x_1 \dots x_n$. Hence, by the Corollary to the Lemma, $G^n_w x_1 \dots x_n$. So by (573.1), $w \models G^n x_1 \dots x_n$.

($\leftarrow$) Assume $w \models G^n x_1 \dots x_n$. Then by (573.1), $G^n_w x_1 \dots x_n$. So by our Lemma, $\Box G^n_w x_1 \dots x_n$, and by a fundamental theorem (543.2), $\forall w'(w' \models G^n_w x_1 \dots x_n)$. Instantiating to w, we have $w \models G^n_w x_1 \dots x_n$.

Case $n = 0$. (Exercise) $\bowtie$

(573.3) [Alternative Proof]. Also by cases $n \geq 1$ and $n = 0$, where n is the arity of G.

Case $n \geq 1$. We know $\exists w \forall p(w \models p \equiv p)$, by (531.1). Let w_1 be such a possible world, so that we know:

(ϑ) $\forall p(w_1 \models p \equiv p)$

Now consider $G^n_{w_1}$. To establish our theorem by showing that $G^n_{w_1}$ is a witness, we have to show, by definition (571.2):

(A) $Rigid(G^n_{w_1})$

(B) $\forall x_1 \dots \forall x_n (G^n_{w_1} x_1 \dots x_n \equiv G x_1 \dots x_n)$

Proof of (A). By (571.1), we have to show:

(C) $G^n_{w_1}\downarrow$

(D) $\Box \forall x_1 \dots \forall x_n (G^n_{w_1} x_1 \dots x_n \rightarrow \Box G^n_{w_1} x_1 \dots x_n)$

Proof of (C). By theorem (570.2).

Proof of (D). By RN and GEN, it suffices to show $G^n_{w_1} x_1 \dots x_n \rightarrow \Box G^n_{w_1} x_1 \dots x_n$. So assume $G^n_{w_1} x_1 \dots x_n$. Then by (573.1), $w_1 \models G^n x_1 \dots x_n$. Hence by (519.1), $\Box w_1 \models G^n x_1 \dots x_n$. But since (573.1) is a modally strict biconditional, it follows by a Rule of Substitution that $\Box G^n_{w_1} x_1 \dots x_n$.

Proof of (B). By GEN, it suffices to show:

$$G^n_{w_1} x_1 \dots x_n \equiv G x_1 \dots x_n$$

But if we substitute $G^n x_1 \dots x_n$ into (ϑ), then we know that:

(ξ) $w_1 \models G^n x_1 \dots x_n \equiv G^n x_1 \dots x_n$

So we may reason as follows:

$$\begin{array}{rcll} G^n_{w_1} x_1 \dots x_n & \equiv & [\lambda x_1 \dots x_n\, w_1 \models G^n x_1 \dots x_n] x_1 \dots x_n & \text{by df (570.1), Rules} =_{df}\text{I}, =_{df}\text{E} \\ & \equiv & w_1 \models G^n x_1 \dots x_n & \text{by } \beta\text{-Conversion} \\ & \equiv & G^n x_1 \dots x_n & \text{by } (\xi) \end{array}$$

$\bowtie$

Case $n = 0$. (Exercise) ⋈

(574.1) We prove this by cases.

Case 1: $n = 0$. Then our theorem asserts $\Box(p \rightarrow \Box p) \equiv (\Diamond p \rightarrow \Box p)$. But this is just an instance of (172.1).

Case 2: $n \geq 1$. By an instance of (172.1), we know:

$$\Box(Fx_1 \ldots x_n \rightarrow \Box Fx_1 \ldots x_n) \equiv (\Diamond Fx_1 \ldots x_n \rightarrow \Box Fx_1 \ldots x_n)$$

So by n applications of GEN:

$$\forall x_1 \ldots \forall x_n(\Box(Fx_1 \ldots x_n \rightarrow \Box Fx_1 \ldots x_n) \equiv (\Diamond Fx_1 \ldots x_n \rightarrow \Box Fx_1 \ldots x_n))$$

So by n applications of (99.3) and a Rule of Substitution:

$$\forall x_1 \ldots \forall x_n \Box(Fx_1 \ldots x_n \rightarrow \Box Fx_1 \ldots x_n) \equiv \forall x_1 \ldots \forall x_n(\Diamond Fx_1 \ldots x_n \rightarrow \Box Fx_1 \ldots x_n)$$

By n applications of BF and the Rule of Substitution (160.2):

$$\Box\forall x_1 \ldots \forall x_n(Fx_1 \ldots x_n \rightarrow \Box Fx_1 \ldots x_n) \equiv \forall x_1 \ldots \forall x_n(\Diamond Fx_1 \ldots x_n \rightarrow \Box Fx_1 \ldots x_n)$$

⋈

(574.2) We prove this by cases.

Case 1: $n = 0$. Then our theorem asserts $\Box(p \rightarrow \Box p) \equiv (\Box p \vee \Box\neg p)$. This is established by the following sequence of equivalences:

$$\begin{array}{lcll}
\Box(p \rightarrow \Box p) & \equiv & \Diamond p \rightarrow \Box p & (172.1) \text{ or } (574.1) \\
 & \equiv & \neg\Diamond p \vee \Box p & (88.1.c) \\
 & \equiv & \Box\neg p \vee \Box p & (162.1), (159.3) \\
 & \equiv & \Box p \vee \Box\neg p & (88.2.c)
\end{array}$$

Case 2: Exercise. ⋈

(574.3) (Exercise)

(577.2) It follows from (490.4) and (487.2) that:

(ϑ) $\forall p(\boldsymbol{s_V} \models p)$

Since $\boldsymbol{s_V}$ is known to be a situation, we have to show: (a) $\boldsymbol{s_V}$ is maximal and (b) $\boldsymbol{s_V}$ fails to be possible. But clearly, if every proposition is true in $\boldsymbol{s_V}$, then for every proposition q, either q is true in $\boldsymbol{s_V}$ or $\neg q$ is true in $\boldsymbol{s_V}$. So (a) holds, by definition of *Maximal* (520). And just as clearly, if every proposition is true in $\boldsymbol{s_V}$, then for every (and hence, for some) proposition q, both q is true in $\boldsymbol{s_V}$

and $\neg q$ is true in $\boldsymbol{s_V}$. So by definition (498), $\neg Consistent(\boldsymbol{s_V})$, and by theorem (504.1), it follows that $\neg Possible(\boldsymbol{s_V})$.[472] ⋈

(577.3) – (577.4) (Exercises)

(578.1) Assume $ImpossibleWorld(x)$. Then by definition (577), we know:

(A) $Situation(x)$

(B) $Maximal(x)$

(C) $\neg Possible(x)$

Our proof strategy is:

- Show (A) $\rightarrow \Box$(A), (B) $\rightarrow \Box$(B), and (C) $\rightarrow \Box$(C)
- Infer from these results that $\Box$((A) & (B) & (C)), by a more general version of (158.3) (exercise)
- Conclude $\Box ImpossibleWorld(x)$, by definition (577) and the Rule of Substitution for Defined Formulas (160.3)

Since the reasoning in the last two steps is straightforward, it remains to show the first. (A) $\rightarrow \Box$(A) holds by the left-to-right direction of (469.1). (B) $\rightarrow \Box$(B) holds, by (522.1). To show (C) $\rightarrow \Box$(C), assume $\neg Possible(x)$. Then it follows, by definition (502) and the Rule of Substitution for Defined Subformulas (160.3), that $\neg\Diamond Actual(x)$. Hence $\Box\neg Actual(x)$, and so by the 4 schema, $\Box\Box\neg Actual(x)$. So by now familiar reasoning, $\Box\neg\Diamond Actual(x)$, and again by definition of $Possible(x)$ and the Rule of Substitution for Defined Subformulas, $\Box\neg Possible(x)$. ⋈

(578.2) – (578.3) From (578.1), by RN and GEN, respectively.

(580) (Exercise)

(581) Our theorem requires us to show:[473]

(a) $Situation(\bot)$

(b) $Maximal(\bot)$

(c) $\neg Possible(\bot)$

---

[472] To see precisely how these arguments for (a) and (b) can be made precise, we can reason as follows. (a) To show maximality, we have to show $\forall p(\boldsymbol{s_V} \models p \vee \boldsymbol{s_V} \models \neg p)$, and so, by GEN, that $\boldsymbol{s_V} \models p \vee \boldsymbol{s_V} \models \neg p$. But by $\forall$E, the first disjunct follows from (ϑ) and by $\vee$I, $\boldsymbol{s_V} \models p \vee \boldsymbol{s_V} \models \neg p$. (b) Let p_1 be any proposition. Then by (104.2), $(p_1 \,\&\, \neg p_1)\downarrow$. So we can instantiate the 0-ary relation term $p_1 \,\&\, \neg p_1$ in (ϑ) and obtain $\boldsymbol{s_V} \models (p_1 \,\&\, \neg p_1)$. But, it is a theorem that $\neg\Diamond(p_1 \,\&\, \neg p_1)$, by (84), RN, and (162.1). Hence by an appropriate instance of (503.2), it follows that $\neg Possible(\boldsymbol{s_V})$.

[473] I'm indebted to Daniel West, who offer a modally strict proof in lieu of the non-modally strict proof originally developed. The following argument follows his proof in many, but not all, details.

(d) $\neg TrivialSituation(\bot)$

(a) is theorem (469.7). We prove (b) – (d) with the aid of two modally strict lemmas:

Lemma 1: $\forall F(\bot F \equiv \mathcal{A}\exists q(\neg q \,\&\, F = [\lambda y\, q]))$
Proof: By definition of $\bot$ (302.2), $\bot = \iota x(A!x \,\&\, \forall F(xF \equiv \exists q(\neg q \,\&\, F = [\lambda y\, q])))$. From this and an appropriate instance of (258.1), if follows that $\bot F \equiv \mathcal{A}\exists q(\neg q \,\&\, F = [\lambda y\, q])$. Since we've derived this from no assumptions, our lemma follows by GEN.

Lemma 2: $\forall p(\bot \models p \equiv \mathcal{A}\neg p)$
Proof: By GEN, it suffices to show $\bot \models p \equiv \mathcal{A}\neg p$. Since we know $\bot$ is a situation, it follows by (471) that $\bot \models p \equiv \bot[\lambda y\, p]$. So by biconditional syllogism, it suffices to show $\bot[\lambda y\, p] \equiv \mathcal{A}\neg p$. ($\rightarrow$) Assume $\bot[\lambda y\, p]$. Then by Lemma 1, $\mathcal{A}\exists q(\neg q \,\&\, [\lambda y\, p] = [\lambda y\, q])$. So by (139.10), $\exists q \mathcal{A}(\neg q \,\&\, [\lambda y\, p] = [\lambda y\, q])$. Assume q_1 is such a proposition, so that we know $\mathcal{A}(\neg q_1 \,\&\, [\lambda y\, p] = [\lambda y\, q_1])$. Hence, by (139.2), we know both $\mathcal{A}\neg q_1$ and $\mathcal{A}[\lambda y\, p] = [\lambda y\, q_1]$. The latter implies $[\lambda y\, p] = [\lambda y\, q_1]$, by (175.1). So by proposition identity, $p = q_1$. Hence $\mathcal{A}\neg p$. ($\leftarrow$) Assume $\mathcal{A}\neg p$. Note that since $[\lambda y\, p] = [\lambda y\, p]$ is a theorem (118.2), it follows by the Rule of Actualization (135) that $\mathcal{A}[\lambda y\, p] = [\lambda y\, p]$. So by &I and (139.2), $\mathcal{A}(\neg p \,\&\, [\lambda y\, p] = [\lambda y\, p])$. Hence $\exists q \mathcal{A}(\neg q \,\&\, [\lambda y\, p] = [\lambda y\, q])$. So by (139.10), $\mathcal{A}\exists q(\neg q \,\&\, [\lambda y\, p] = [\lambda y\, q])$. Then by Lemma 1, $\bot[\lambda y\, p]$.

We then establish (b) – (d) as follows.

(b) By GEN and the definition of maximality, we have to show $\bot \models p \vee \bot \models \neg p$. We reason to this last claim by disjunctive syllogism from theorem (139.1), the *commuted* form of which tells us $\mathcal{A}\neg p \vee \mathcal{A}p$. Assume $\mathcal{A}\neg p$. Then by Lemma 2, $\bot \models p$. Assume $\mathcal{A}p$. Since it is a modally strict theorem that $p \equiv \neg\neg p$, it follows by a Rule of Substitution that $\mathcal{A}\neg\neg p$. Hence, by Lemma 2, $\bot \models \neg p$. So by disjunctive syllogism (86.3.c), $\bot \models p \vee \bot \models \neg p$.

(c) Consider any proposition q. Since $\neg(q \,\&\, \neg q)$ is a theorem it follows by the Rule of Actualization that $\mathcal{A}\neg(q \,\&\, \neg q)$. Hence, by Lemma 2, $\bot \models (q \,\&\, \neg q)$. Since $\neg(q \,\&\, \neg q)$ is also a modally strict theorem, it follows by RN that $\Box\neg(q \,\&\, \neg q)$ and, by modal negation, $\neg\Diamond(q \,\&\, \neg q)$. So by &I and existentially generalizing on the proposition $q \,\&\, \neg q$, we have established $\exists p((\bot \models p) \,\&\, \neg\Diamond p)$. Since $\bot$ is a situation, it follows by (503.2) that $\neg Possible(\bot)$.

(d) By definition (487.2), to show that $\bot$ isn't a trivial situation, we have to show that either it isn't a situation or that there is some proposition that it doesn't make true. Since we know $\bot$ is a situation, we therefore have to show $\exists q\neg(\bot \models q)$. As our witness, consider any proposition that we can prove by modally strict means, say, p_0, which was defined in (208) as $\forall x(E!x \rightarrow E!x)$. We want to show $\neg\bot \models p_0$. Assume, for reductio, that $\bot \models p_0$. Then by Lemma 2, $\mathcal{A}\neg p_0$, and so by axiom (44.1), $\neg\mathcal{A}p_0$. But since p_0 is provable by modally strict means, it follows by RN that $\Box p_0$ and, hence, that $\mathcal{A}p_0$, by (132). Contradiction. $\bowtie$

(582) We prove this theorem with the help of one of the lemmas (Lemma 2) established in the proof of (581):

> *Lemma 2*: $\forall p(\bot \models p \equiv \mathcal{A}\neg p)$

By definition of 1-modal closure (527), we have to show that there exist propositions p and q such that $\bot \models p$, $p \Rightarrow q$, and $\neg(\bot \models q)$. Let our witness for p be $p_0 \,\&\, \neg p_0$, where p_0 is $\forall x(E!x \to E!x)$, and let our witness for q be p_0. So we have to show:

(i) $\bot \models (p_0 \,\&\, \neg p_0)$

(ii) $(p_0 \,\&\, \neg p_0) \Rightarrow p_0$

(iii) $\neg(\bot \models p_0)$

(i) Since $p_0 \,\&\, \neg p_0$ is a contradiction, we know $\neg(p_0 \,\&\, \neg p_0)$ is a theorem (84). Hence, by Rule RA, $\mathcal{A}\neg(p_0 \,\&\, \neg p_0)$. So by Lemma 2, $\bot \models (p_0 \,\&\, \neg p_0)$.

(ii) By (85.1), $(p_0 \,\&\, \neg p_0) \to p_0$. Since this is a modally strict theorem, it follows by RN that $\Box((p_0 \,\&\, \neg p_0) \to p_0)$. So by definition of $\Rightarrow$ (442.1), it follows that $(p_0 \,\&\, \neg p_0) \Rightarrow p_0$.

(iii) $\neg(\bot \models p_0)$ was established in part (d) of the proof of (581). $\bowtie$

(584.1) Let φ be formula $s \models q \vee q = p$. By GEN, it suffices to show, by modally strict means, that $\varphi \to \Box\varphi$. So assume φ, i.e.,

$$s \models q \vee q = p$$

Note that both disjuncts imply their own necessity: $s \models q$ implies $\Box s \models q$ by (473.1), and $q = p$ implies $\Box q = p$ by (125.1). Hence, by disjunctive syllogism:

$$\Box s \models q \vee \Box q = p$$

But by (158.15), it follows that $\Box(s \models q \vee q = p)$. $\bowtie$

(584.2) (Exercise)

(584.3) Assume $s \models q$. Hence by $\vee$I, $s \models q \vee q = p$. So by (584.2), it follows that $s^{+p} \models q$. $\bowtie$

(584.4) (Exercise)

(585) By theorem (531.1), we know $\exists w \forall q(w \models q \equiv q)$. Let w_0 be an arbitrary such possible world, so that we know:

(ϑ) $\forall q(w_0 \models q \equiv q)$

Note that, by definition, w_0 is a situation. Now to establish the present theorem by conditional proof, assume $\neg\Diamond p$. Then consider the p-extension of w_0, i.e., ${w_0}^{+p}$, as this is defined in (583). Clearly, ${w_0}^{+p}$ is a situation (exercise). Since both w_0 and ${w_0}^{+p}$ are situations, it follows from (584.4) that ${w_0}^{+p} \models p$. So to show that ${w_0}^{+p}$ is a witness to $\exists i(\neg \mathit{TrivialSituation}(i) \,\&\, i \models p)$, it remains, by the definition of an impossible world (577.1), &I, and $\exists$I, to show:

(a) $\mathit{Maximal}({w_0}^{+p})$

(b) $\neg \mathit{Possible}({w_0}^{+p})$

(c) $\neg \mathit{TrivialSituation}({w_0}^{+p})$

(a) Since w_0 is a possible world, we know that it is maximal (521). But by (584.3) and the fact that possible worlds are situations, we know that every proposition true at w_0 is true at ${w_0}^{+p}$. Hence ${w_0}^{+p}$ is maximal, by disjunctive syllogism from the maximality of w_0.

(b) It follows from our initial assumption that $\Box\neg p$. So $\forall w(w \models \neg p)$, by a fundamental theorem (543.2) of possible world theory. Hence $w_0 \models \neg p$. So by (584.3) and the fact that possible worlds are situations, ${w_0}^{+p} \models \neg p$. But we already know ${w_0}^{+p} \models p$. By conjoining our last two results in reverse order and generalizing, we have $\exists q({w_0}^{+p} \models q \,\&\, {w_0}^{+p} \models \neg q)$. Hence, by definition, $\neg \mathit{Consistent}({w_0}^{+p})$. Thus, by (504.1), $\neg \mathit{Possible}({w_0}^{+p})$.

(c) To show that $\neg \mathit{TrivialSituation}({w_0}^{+p})$, we have to show $\exists q \neg({w_0}^{+p} \models q)$, i.e., find a proposition that isn't true at ${w_0}^{+p}$.[474] If we choose our witness to be a contingent falsehood, then we can show that such a proposition is provably distinct from p and fails to be true in ${w_0}^{+p}$, as follows. By (217.2), we know that there are contingently false propositions. So let q_1 be such a proposition. Since q_1 is false, it follows from (ϑ) that $\neg w_0 \models q_1$. Moreover, since q_1 is contingently false and p is impossible, it follows by (214.6) that $q_1 \neq p$. Hence $\neg(w_0 \models q_1 \vee q_1 = p)$. So by (584.2), $\neg({w_0}^{+p} \models q_1)$. Hence $\exists q \neg({w_0}^{+p} \models q)$. $\bowtie$

(586.1) By (581), $\perp$ is an impossible world. In (582), we established that $\perp$ is not 1-modally closed. In the proof of the latter, we saw that when p_0 is the necessary truth $\forall x(E!x \to E!x)$, then the contradiction $p_0 \,\&\, \neg p_0$ is true at $\perp$ but that p_0 fails to be true at $\perp$, i.e., that both $\perp \models (p_0 \,\&\, \neg p_0)$ and $\neg(\perp \models p_0)$. Hence,

[474] The witness can't be a true proposition, for by (ϑ), all and only true propositions are true at w_0 and so the propositions true at w_0 are true at ${w_0}^{+p}$ (584.3). Nor can the witness be an arbitrary falsehood, since the arbitrary choice might be p itself, which is true in ${w_0}^{+p}$ despite being necessarily false. The witness can't be some necessary falsehood distinct from p because we haven't established that there is such a falsehood. Though our system allows us to assert, without contradiction, that there are distinct necessary falsehoods, our axioms thus far don't guarantee that there are such. See the discussion in footnote 306.

there is an impossible world i and propositions p $(= p_0)$ and q $(= p_0)$ such that $p \,\&\, \neg p$ is true at i but q fails to be true at i. ⋈

(586.2) Again let p_0 be the proposition $\forall x(E!x \to E!x)$ and consider $\overline{p_0}$, which by (199.7) is defined as $\neg\forall x(E!x \to E!x)$. Clearly, $\neg\Diamond\overline{p_0}$. Now theorem (531.1) tells us that $\exists w \forall p(w \models p \equiv p)$. Let w_0 be such a world, so that we know $\forall p(w_0 \models p \equiv p)$. Consider the $\overline{p_0}$-extension of w_0, namely, ${w_0}^{+\overline{p_0}}$. By reasoning analogous to that used in the proof of (585), it follows that (exercise):

- $ImpossibleWorld({w_0}^{+\overline{p_0}})$
- ${w_0}^{+\overline{p_0}} \models \overline{p_0}$
- ${w_0}^{+\overline{p_0}} \models \neg\overline{p_0}$ (since $\neg\overline{p_0}$ is true)

Moreover, by reasoning analogous to part (c) of the proof of (585), there is a contingently false proposition, say q_1, that fails to be true at ${w_0}^{+\overline{p_0}}$. Hence, there is an impossible world i and there are propositions p $(= \overline{p_0})$ and q $(= q_1)$ such that p and $\neg p$ are both true at i while q fails to be true at i. ⋈

(587) By (531.1), we know that there is a possible world that makes true all and only the true propositions. Let w_0 be such a world, so that we know:

(ϑ) $\forall p(w_0 \models p \equiv p)$

Let p_0 be $\forall x(E!x \to E!x)$, so that $\overline{p_0}$ is the negation of p_0. Now consider ${w_0}^{+\overline{p_0}}$, i.e., w_0 extended with necessary falsehood $\overline{p_0}$. So by reasoning analogous to that in (585), it can be established that ${w_0}^{+\overline{p_0}}$ is an impossible world (exercise). Finally, by theorem (217.2), there is a contingently false proposition, say r_0. With these facts in hand, it suffices by &I and $\exists$I to prove the following to establish our theorem:

(a) ${w_0}^{+\overline{p_0}} \models (p_0 \vee r_0)$

(b) ${w_0}^{+\overline{p_0}} \models \neg p_0$

(c) $\neg {w_0}^{+\overline{p_0}} \models r_0$

To prove these claims, we shall, on occasion, rehearse some of the steps in the proof of (585).

(a) Since p_0 is a necessary truth, it follows by a Fundamental Theorem of Possible World Theory (543.2) that $\forall w(w \models p_0)$. Hence $w_0 \models p_0$. Since possible worlds are 1-modally closed (528), it follows that $w_0 \models (p_0 \vee r_0)$. But since w_0 is a situation, we know, by (584.3), that every proposition true at w_0 is true in ${w_0}^{+\overline{p_0}}$. Hence ${w_0}^{+\overline{p_0}} \models (p_0 \vee r_0)$.

(b) As as instance of (586.3), we know ${w_0}^{+\overline{p_0}} \models \overline{p_0}$. But by theorem (199.7), $\overline{p_0} = \neg p_0$. Hence, ${w_0}^{+\overline{p_0}} \models \neg p_0$.

(c) The proof of $\neg {w_0}^{+\overline{p_0}} \models r_0$ involves reasoning analogous to part (c) of the proof of (585) and will appeal to the following instance of (584.2):

$$\forall q({w_0}^{+\overline{p_0}} \models q \equiv w_0 \models q \vee q = \overline{p_0})$$

The cases are analogous because: (1) ${w_0}^p$ in (585) and ${w_0}^{+\overline{p_0}}$ in the present theorem are both impossible worlds constructed by extending w_0 with a necessary falsehood, and (2) in part (c) of the proof of (585), it was established that in such impossible worlds, contingently false propositions fail to be true. So we leave the rest of the proof as an exercise. ⋈

(592.2) (Exercise)

(596) (Exercise)

(597) Since $\underline{s}$ is a weak restricted variable, it is best to eliminate it to avoid any modal reasoning errors. Once we eliminate the restricted variable, we have to show:

$$Story(x) \to (x = \imath s \forall p(s \models p \equiv x \models p))$$

So assume $Story(x)$. (Although this assumption is modally fragile, it will be discharged by our conditional proof, so that the resulting theorem is derived by modally strict means.) Then by the definition of a story (592.1), we know $Situation(x)$. This means we're now trying to establish the identity of situations, and so it suffices by (474) to show that they make the same propositions true propertie, i.e., for an arbitrary proposition q, we have to show that:

$$(x \models q) \equiv \imath s \forall p(s \models p \equiv x \models p) \models q$$

Note that if we apply GEN to theorem (486.4), we obtain:

$$\forall y(y = \imath s \forall p(s \models p \equiv \varphi) \to \forall p(y \models p \equiv \varphi)),$$
provided φ is a rigid condition on propositions

Now since the description $\imath s \forall p(s \models p \equiv \varphi)$ is canonical and hence significant, we may instantiate it into the above universal claim to obtain:

$$\imath s \forall p(s \models p \equiv \varphi) = \imath s \forall p(s \models p \equiv \varphi) \to \forall p(\imath s \forall p(s \models p \equiv \varphi) \models p \equiv \varphi),$$
provided φ is a rigid condition on propositions

It also follows from the fact that the description $\imath s \forall p(s \models p \equiv x \models p)$ is signficant that the antecedent to the above claim holds. Hence:

(ϑ) $\forall p(\imath s \forall p(s \models p \equiv \varphi) \models p \equiv \varphi)$,
provided φ is a rigid condition on propositions

Now consider the formula $x \models p$. Since x is a story, then the fact that (596) is a modally strict theorem implies that $x \models p$ is a rigid condition on propositions. Moreover, s is not free in the formula $x \models p$. So, if we take φ in (ϑ) to be $x \models p$, the resulting instance of (ϑ) is:

$$\forall p(\imath s\forall p(s \models p \equiv x \models p) \models p \equiv (x \models p))$$

Now if we instantiate our arbitrary proposition q into this universal claim, we obtain:

$$\imath s\forall p(s \models p \equiv x \models p) \models q \equiv (x \models q)$$

By the commutativity of the biconditional, we have obtained what we had to show, namely:

$$(x \models q) \equiv \imath s\forall p(s \models p \equiv x \models p) \models q \qquad \bowtie$$

(603.1) Assume $A!x$. Then by (222.3), $\neg O!x$. So by (115.3), it follows that $\neg\Diamond E!x$. But as an instance of (126.1), we know that the following is a modally strict theorem:

$$\Diamond E!x \equiv \exists y(y = x \,\&\, \Diamond E!y)$$

By the commutativity of &, so is:

$$\Diamond E!x \equiv \exists y(\Diamond E!y \,\&\, y = x)$$

Hence, by a Rule of Substitution, $\neg\exists y(y = x \,\&\, \Diamond E!y)$. $\bowtie$

(603.2) – **(603.3)** (Exercises)

(607) (Exercises)

(608.1) In the following derivation, let s denote *A Study in Scarlet*, h denote Sherlock Holmes, and D denote *being a detective*. Assume both $s \models Dh_s$ and *OriginalCharacterOf*(h_s, s) as premises. The second implies, by axiom (600), that:

(ϑ) $h_s = \imath x(A!x \,\&\, \forall F(xF \equiv s \models Fh_s))$

By theorem (596), we know the formula $s \models Fh_s$ is a rigid condition on properties. Hence (ϑ) and (261.2) imply $A!h_s$ and $\forall F(h_sF \equiv s \models Fh_s)$. So by our first premise, it follows that h_sD. $\bowtie$

(608.2) In the following derivation, let i denote *The Iliad*, F denote the relation *x fought y*, a denote Achilles and h denote Hector. Then we have, as assumptions:

(A) $i \models Fa_ih_i$

(B) $OriginalCharacterOf(a_i, i)$

(C) $OriginalCharacterOf(h_i, i)$

Since our background assumption is that β-Conversion holds under the scope of the story operator, (A) implies both:

(D) $i \models [\lambda x\, Fxh_i]a_i$

(E) $i \models [\lambda x\, Fa_ix]h_i$

Hence by the reasoning in (608.1), it follows both at $a_i[\lambda x Fxh_i]$ and $h_i[\lambda x Fa_ix]$. But, as an instance of axiom (50), we know:

$$a_ih_iF \equiv (a_i[\lambda x\, Fxh_i]\ \&\ h_i[\lambda x\, Fa_ix])$$

Hence, it follows that a_ih_iF. ⋈

(613) (Exercise)

(614.1) – **(614.4)** (Exercises)

(616.1) – **(616.3)** (Exercises)

(618.1) – **(618.3)** (Exercises)

(620.1) (Exercise)

(620.2) By GEN, it suffices to show $(dF \vee eF) \to \Box(dF \vee eF)$. So assume $dF \vee eF$. Since axiom (51) applies to every individual whatsoever, it applies to concepts and so each disjunct of our assumption implies its own necessitation. Hence, by disjunctive syllogism, $\Box dF \vee \Box eF$. By (158.15), it follows that $\Box(dF \vee eF)$. ⋈

(620.3) Let φ be the formula $dF \vee eF$. Then by (620.2), it is a rigid condition on properties. So by (261.2), we know:

$$y = \imath x(A!x\ \&\ \forall F(xF \equiv dF \vee eF)) \to (A!y\ \&\ \forall F(yF \equiv dF \vee eF))$$

From the identity (614.4), it follows by Rule =E that:

$$y = \imath x(C!x\ \&\ \forall F(xF \equiv dF \vee eF)) \to (A!y\ \&\ \forall F(yF \equiv dF \vee eF))$$

Moreover, by definition of $C!$ (612), it follows from this last result by Rule =E that:

$$y = \imath x(C!x\ \&\ \forall F(xF \equiv dF \vee eF)) \to (C!y\ \&\ \forall F(yF \equiv dF \vee eF))$$

Furthermore, by our conventions for restricted variables, this previous result can be expressed as:

$$y = \imath c\forall F(cF \equiv dF \vee eF) \to (C!y\ \&\ \forall F(yF \equiv dF \vee eF))$$

So by GEN, the fact that $(d \oplus e)\downarrow$, and Rule =E, it follows that:

$$d \oplus e = \imath c \forall F(cF \equiv dF \vee eF) \rightarrow C!d \oplus e \;\&\; \forall F(d \oplus eF \equiv dF \vee eF)$$

Hence, by definition of $d \oplus e$ (619) it follows that:

$$C!d \oplus e \;\&\; \forall F(d \oplus eF \equiv dF \vee eF) \qquad \bowtie$$

(620.4) (Exercise)

(621.1) Given that $(c \oplus c)\downarrow$, that $c \oplus c$ and c are both concepts (620.3), and that both are abstract (612), it suffices by (245.2) to show $c \oplus c$ and c encode the same properties. So, by GEN, we show $c \oplus cF \equiv cF$, as follows:

$$\begin{array}{rcll} c \oplus cF & \equiv & cF \vee cF & \text{by (620.3)} \\ & \equiv & cF & \text{by idempotence of } \vee \text{ (85.7)} \qquad \bowtie \end{array}$$

(621.2) (Exercise)

(621.3) [In the following proof, we sometimes assert unary encoding formulas of the form $\kappa\Pi$ in which κ is either $(c \oplus d) \oplus e$ or $c \oplus (d \oplus e)$ and Π is F.] Since $((c \oplus d) \oplus e)\downarrow$ and $(c \oplus (d \oplus e))\downarrow$, and both $(c \oplus d) \oplus e$ and $c \oplus (d \oplus e)$ are abstract objects, it suffices by (245.2) to show they encode the same properties. So, by GEN, we show $(c \oplus d) \oplus eF \equiv c \oplus (d \oplus e)F$ as follows:

$$\begin{array}{rcll} (c \oplus d) \oplus eF & \equiv & c \oplus dF \vee eF & \text{by (620.3)} \\ & \equiv & (cF \vee dF) \vee eF & \text{by (620.3) and (88.8.h)} \\ & \equiv & cF \vee (dF \vee eF) & \text{by associativity of } \vee \text{ (88.2.d)} \\ & \equiv & cF \vee d \oplus eF & \text{by (620.3) and (88.8.g)} \\ & \equiv & c \oplus (d \oplus e)F & \text{by (620.3)} \qquad \bowtie \end{array}$$

(623.1) – **(623.2)** (Exercises)

(625.1) (Exercise)

(625.2) Suppose $c \preceq d$ and $c \neq d$. To show that $d \not\preceq c$, we must establish that there is a property that d encodes but which c doesn't encode. Now since c and d are both concepts, they are both abstract objects. Since they are distinct, it follows by (245.3) that either there is a property c encodes that d doesn't, or there is a property d encodes that c doesn't. But, since $c \preceq d$ it follows by definition of $\preceq$ (624.1) that it must be the latter. $\bowtie$

(625.3) (Exercise)

(626.1) (Exercise)

(626.2) ($\rightarrow$) Exercise. ($\leftarrow$) Assume $\forall e(e \preceq c \equiv e \preceq d)$, to show that $c = d$. Since concepts are abstract objects, it suffices by GEN to show $cF \equiv dF$:

(→) Assume cF and, for reductio, assume $\neg dF$. So $c \not\preceq d$. But it follows from our initial hypothesis that $c \preceq c \equiv c \preceq d$. Hence $c \not\preceq c$, which contradicts the reflexivity of concept inclusion (625.1).

(←) By analogous reasoning. ⋈

(626.3) (→) Exercise. (←) Assume $\forall e(c \preceq e \equiv d \preceq e)$. By now familiar reasoning, it suffices to show that $cF \equiv dF$:

(→) Assume cF and, for reductio, assume $\neg dF$. So $c \not\preceq d$. But it follows from our initial hypothesis that $c \preceq d \equiv d \preceq d$. Hence $d \not\preceq d$, which contradicts the reflexivity of concept inclusion (625.1).

(←) By analogous reasoning. ⋈

(627.1) – **(627.3)** (Exercises)

(627.4) (→) Assume $c \oplus d \preceq e$. To show $c \preceq e$, we have to show $\forall F(cF \rightarrow eF)$. So assume cF. Then $cF \vee dF$. So by the second conjunct of the modally-strict theorem (620.3) about sums, it follows that $c \oplus dF$. But by definition of $\preceq$, our initial assumption implies $\forall F(c \oplus dF \rightarrow eF)$. Hence eF. One can show $d \preceq e$ by analogous reasoning.

(←) Assume $c \preceq e \,\&\, d \preceq e$. To show $c \oplus d \preceq e$, it suffices by GEN to show: $c \oplus dF \rightarrow eF$. Assume $c \oplus dF$. Then by the second conjunct of the modally-strict theorem (620.3) about sums, it follows that $cF \vee dF$. Reasoning by cases: if cF, then by the first conjunct of our initial assumption, it follows that eF, and if dF, then by the second conjunct of our initial assumption, it follows that eF. Hence, eF. ⋈

(627.5) (Exercise)

(628) (→) Assume $c \preceq d$. We prove $\exists e(c \oplus e = d)$ by cases, with the two cases being: (a) $c = d$ and (b) $c \neq d$. (a) Suppose $c = d$. By the idempotency of $\oplus$, $c \oplus c = c$, in which case, $c \oplus c = d$. Therefore, $\exists e(c \oplus e = d)$. (b) Suppose $c \neq d$. Then since $c \preceq d$, it follows that $d \not\preceq c$ (625.2), and so we know there must be one or more properties encoded by d which are not encoded by c. By Comprehension for Concepts (614.1), we know there exists a concept that encodes those properties F that d encodes and c doesn't:

$$\exists c' \forall F(c'F \equiv dF \,\&\, \neg cF)$$

Let c_1 be an arbitrary such concept, so that we know:

$$(\vartheta) \quad \forall F(c_1 F \equiv dF \,\&\, \neg cF)$$

To complete the proof of (b), it suffices by $\exists$I to show $c \oplus c_1 = d$, i.e., that $c \oplus c_1$ and d encode the same properties:

(→) Assume $c \oplus c_1 G$ (to show: dG). By (620.3), it follows that $cG \vee c_1 G$. Reasoning by cases: if cG, then by the fact that $c \leq d$, it follows that dG; if $c_1 G$, then by (ϑ), it follows that $dG \mathbin{\&} \neg cG$, and hence dG.

(←) Assume dG (to show $c \oplus c_1 G$). This time our proof by cases begins from $cG \vee \neg cG$. If cG, then $cG \vee c_1 G$, so by (620.3), $c \oplus c_1 G$. Alternatively, if $\neg cG$, then we have $dG \mathbin{\&} \neg cG$. So by (ϑ), $c_1 G$, and hence $cG \vee c_1 G$. So by (620.3), it follows that $c \oplus c_1 G$.

(←) Assume $\exists e(c \oplus e = d)$. Let c_2 be an arbitrary such concept, so that we know $c \oplus c_2 = d$. To show $c \leq d$, assume cG (to show dG). Then, $cG \vee c_2 G$, which by (620.3) entails that $c \oplus c_2 G$. But by hypothesis, $c \oplus c_2 = d$. So dG. ⋈

(629) (→) Assume $c \leq d$. So $\forall F(cF \rightarrow dF)$. To show that $c \oplus d = d$, it suffices, by now familiar reasoning, to show that $c \oplus d$ and d encode the same properties:

(→) Assume $c \oplus dG$. Then, by (620.3), $cG \vee dG$. If we reason by cases from the two disjuncts to the conclusion dG, then it suffices to show $cG \rightarrow dG$. But the assumption that cG imples dG, by the fact that $c \leq d$.

(←) Assume dG. Then $cG \vee dG$. So by (620.3), $c \oplus dG$.

(←) Assume that $c \oplus d = d$. It follows that $\exists e(c \oplus e = d)$. So by (628), $c \leq d$. ⋈

(630.1) Assume $(c \npreceq d) \mathbin{\&} (d \npreceq c))$. We show that $c \oplus d$ is a witness to $\exists e(e \neq c \mathbin{\&} e \neq d \mathbin{\&} c \oplus e = c \oplus d)$. We leave it as an exercise to show that $c \oplus d \neq c$ and $c \oplus d \neq d$. To show $c \oplus (c \oplus d) = c \oplus d$, we may reason as follows:

$$\begin{array}{rcll} c \oplus (c \oplus d) & = & c \oplus (c \oplus d) & \text{Rule =I} \\ & = & (c \oplus c) \oplus d & \text{by associativity (621.3)} \\ & = & c \oplus d & \text{by idempotence (621.1)} \qquad ⋈ \end{array}$$

(630.2) (→) Assume $c \leq d$ and $d \npreceq c$. By Comprehension for Concepts (614.1), we know:

$\exists c' \forall F(c'F \equiv dF \mathbin{\&} \neg cF)$

Let e_1 be such a concept, so that we know:

(ϑ) $\forall F(e_1 F \equiv dF \mathbin{\&} \neg cF)$

Now by &I and ∃I, we want to show (a) $e_1 \npreceq c$, and (b) $c \oplus e_1 = d$:

(a) Since $d \npreceq c$ by assumption, $\exists F(dF \mathbin{\&} \neg cF)$. Let P be an arbitrary such property, so that we know $dP \mathbin{\&} \neg cP$. Then by (ϑ), it follows that $e_1 P$. Hence, we've established $e_1 P \mathbin{\&} \neg cP$. So $e_1 \npreceq c$.

(b) Since concepts are abstract objects, it suffices by (245.2) and GEN to show $c \oplus e_1 G \equiv dG$. (→) Assume $c \oplus e_1 G$. Then $cG \vee e_1 G$. So we may reason by cases. If cG, then since $c \preceq d$, dG. If $e_1 G$, then by (ϑ), it follows that $dG \,\&\, \neg cG$. *A fortiori*, dG. (←) Assume dG. Now we reason by cases from the tautology $cG \vee \neg cG$. If cG, then $cG \vee e_1 G$, and hence $c \oplus e_1 G$. If $\neg cG$, then $dG \& \neg cG$, and hence, by (ϑ), $e_1 G$. Therefore, $cG \vee e_1 G$, and so $c \oplus e_1 G$.

(←) Assume $\exists e(e \npreceq c \,\&\, c \oplus e = d)$. Let e_2 be such a concept, so that we know:

(ξ) $e_2 \npreceq c \,\&\, c \oplus e_2 = d$

The right conjunct of (ξ) implies $\exists e(c \oplus e = d)$. So by (628), $c \preceq d$. Hence it remains to show $d \npreceq c$, i.e., that $\exists F(dF \,\&\, \neg cF)$. Note that the first conjunct of (ξ) implies $\exists F(e_2 F \,\&\, \neg cF)$. Suppose Q is such a property, so that we know both $e_2 Q$ and $\neg cQ$. From the former, it follows that $cQ \vee e_2 Q$ and hence that $c \oplus e_2 Q$, by the second conjunct of (620.3). So by the second conjunct of (ξ) and Rule =E, it follows that dQ. So we have established $dQ \,\&\, \neg cQ$. Hence $\exists F(dF \,\&\, \neg cF)$. ⋈

(632.1) – **(632.3)** (Exercises)

(634.1) – **(634.4)** (Exercises)

(635.1) – **(635.3)** (Exercises)

(636.1) It suffices to show $c \oplus (c \otimes d)F \equiv cF$. (→) Assume $c \oplus (c \otimes d)F$. Then by (620.3), we know $cF \vee c \otimes dF$. Reasoning by cases, it remains only to show $c \otimes dF \rightarrow cF$. So assume $c \otimes dF$. Then by (634.3), both $cF \,\&\, dF$. *A fortiori*, cF. (←) Assume cF. Then $cF \vee c \otimes dF$. Hence, by (620.3), $c \oplus (c \otimes d)F$. ⋈

(636.2) (Exercise)

(637.1) It suffices to show that $c \oplus \boldsymbol{a}_\varnothing F \equiv cF$. (→) Assume $c \oplus \boldsymbol{a}_\varnothing F$. Then, by theorem (620.3), $cF \vee \boldsymbol{a}_\varnothing F$. But by theorem (266.3) and definition (263.1), we know $\boldsymbol{a}_\varnothing$ doesn't encode any properties and so $\neg \boldsymbol{a}_\varnothing F$. Hence cF. (←) Assume cF. Then $cF \vee \boldsymbol{a}_\varnothing F$. Hence, by theorem (620.3), $c \oplus \boldsymbol{a}_\varnothing F$. ⋈

(637.2) It suffices to show that $c \otimes \boldsymbol{a}_{\boldsymbol{V}} F \equiv cF$. (→) Assume $c \otimes \boldsymbol{a}_{\boldsymbol{V}} F$. Then, by theorem (634.3), $cF \,\&\, \boldsymbol{a}_{\boldsymbol{V}} F$. *A fortiori*, cF. (←) Assume cF. But by theorem (266.4) and definition (263.2), we know $\boldsymbol{a}_{\boldsymbol{V}}$ encodes every property. So $\boldsymbol{a}_{\boldsymbol{V}} F$. Hence $cF \,\&\, \boldsymbol{a}_{\boldsymbol{V}} F$. So by (634.3), $c \otimes \boldsymbol{a}_{\boldsymbol{V}} F$. ⋈

(637.3) – **(637.4)** (Exercises)

(639.1) It suffices to show that $c \oplus (d \otimes e)F \equiv (c \oplus d) \otimes (c \oplus e)F$:

$$\begin{array}{rcll}
c{\oplus}(d{\otimes}e)F & \equiv & cF \vee d{\otimes}eF & \text{by (620.3)}\\
& \equiv & cF \vee (dF \,\&\, eF) & \text{by (634.3) and (88.8.g)}\\
& \equiv & (cF \vee dF) \,\&\, (cF \vee eF) & \text{by (88.6.b)}\\
& \equiv & (c{\oplus}dF) \,\&\, (cF \vee eF) & \text{by (620.3) and (88.4.e)}\\
& \equiv & (c{\oplus}dF) \,\&\, (c{\oplus}eF) & \text{by (620.3) and (88.4.f)}\\
& \equiv & (c{\oplus}d){\otimes}(c{\oplus}e)F & \text{by (634.3)} \qquad \bowtie
\end{array}$$

(639.2) (Exercise)

(641.1) – **(641.3)** (Exercises)

(643.1) (Exercise)

(643.2) By GEN and RN, it suffices to show $\neg dF \rightarrow \Box\neg dF$. But as an instance of theorem (179.7), we know $\neg dF \equiv \Box\neg dF$. *A fortiori*, $\neg dF \rightarrow \Box\neg dF$. ⋈

(643.3) – **(643.4)** (Exercises)

(644.1) Theorem (266.4) is that $\mathit{Universal}(\boldsymbol{a_V})$. By definition (263.2), it follows that $\forall F(\boldsymbol{a_V}F)$. So to show $c \oplus {-c}$ and $\boldsymbol{a_V}$ are identical, i.e., encode the same properties, it suffices to show that $\forall F(c \oplus {-cF})$ and, by GEN, that $c \oplus {-cF}$. Now it is a tautology that $cF \vee \neg cF$. Since $\neg cF \equiv -cF$ is a consequence of the second conjunct of (643.3), it follows by (88.8.g) that $cF \vee -cF$. But from this it follows by (620.3) that $c \oplus {-cF}$. ⋈

(644.2) (Exercise)

(645.1) It suffices to show $--c$ and c encode the same properties. Before we prove this, note that by the second conjunct of (643.3), we know:

(ϑ) $\forall F(-cF \equiv \neg cF)$

Since it is a modally strict theorem that a biconditional is equivalent to the result of negating both sides, we know $(-cF \equiv \neg cF) \equiv (\neg{-cF} \equiv \neg\neg cF)$. So by the Rule of Substitution (160.2), it follows from (ϑ) that:

(ξ) $\forall F(\neg{-cF} \equiv \neg\neg cF)$

So we may prove our theorem as follows:

$$\begin{array}{rcll}
--cF & \equiv & \neg{-cF} & \text{by (643.3), } -c \text{ substituted for } d\\
& \equiv & \neg\neg cF & \text{by } (\xi)\\
& \equiv & cF & \text{by } \neg\neg\text{E (78.2)} \qquad \bowtie
\end{array}$$

(645.2) It suffices to show $-c \oplus -d$ encodes F if and only if $-(c \otimes d)$ encodes F:

$$\begin{aligned} -c \oplus -dF &\equiv -cF \vee -dF && \text{by (620.3)} \\ &\equiv \neg cF \vee -dF && \text{by (643.3) and (88.8.h)} \\ &\equiv \neg cF \vee \neg dF && \text{by (643.3) and (88.8.g)} \\ &\equiv \neg(cF \,\&\, dF) && \text{by (88.5.c)} \\ &\equiv \neg c \otimes dF && \text{by (634.3) and (88.4.b)} \\ &\equiv -(c \otimes d)F && \text{by (643.3)} \end{aligned}$$ ⋈

(645.3) (Exercise)

(650.1) Assume for reductio that $c \prec c$. Then by definition of $\prec$ (649), it follows that $c \preceq c$ and $c \neq c$. But by Rule =I, $c = c$. Contradiction. ⋈

(650.2) Assume $c \prec d$ and $d \prec e$. By (649), it follows from these two assumptions, respectively, that:

(ϑ) $c \preceq d \,\&\, c \neq d$

(ξ) $d \preceq e \,\&\, d \neq e$

Now to show $c \prec e$, we have to show $c \preceq e \& c \neq e$, by (649). The first conjuncts of (ϑ) and (ξ) jointly imply $c \preceq e$, by the transitivity of $\preceq$ (625.3). So it remains to show $c \neq e$. Assume, for reductio, that $c = e$. Then substituting e for c into the first conjunct of (ϑ), it follows that $e \preceq d$. But from this and the first conjunct of (ξ), it follows that $e = d$, by the right-to-left direction of (626.1). But by the symmetry of identity, this contradicts the second conjunct of (ξ). ⋈

(650.3) Assume $c \prec d$. For reductio, assume $d \prec c$. Then by (650.2), $c \prec c$, contradicting (650.1). ⋈

(652.1) By definition (651.1), we have to show $\forall d(\boldsymbol{a}_{\varnothing} \preceq d)$. It suffices, by GEN, to show $\boldsymbol{a}_{\varnothing} \preceq d$, and so by definition (624.1), we must show $\forall F(\boldsymbol{a}_{\varnothing}F \to dF)$. Again, by GEN, we show $\boldsymbol{a}_{\varnothing}F \to dF$. But independently, by theorem (266.3) and definition (263.1), we know $\neg\exists F\boldsymbol{a}_{\varnothing}F$, i.e., $\forall F\neg\boldsymbol{a}_{\varnothing}F$. Instantiating to F, it follows that $\neg\boldsymbol{a}_{\varnothing}F$. Hence $\boldsymbol{a}_{\varnothing}F \to dF$, by failure of the antecedent. ⋈

(652.2) Since we know $Bottom(\boldsymbol{a}_{\varnothing})$, it remains only to show $\forall c(Bottom(c) \to c = \boldsymbol{a}_{\varnothing})$, since our theorem will then follow by &I, ∃I, and the definition of the unique existence quantifier. By GEN, it suffices to show $Bottom(c) \to c = \boldsymbol{a}_{\varnothing}$. So assume $Bottom(c)$. Since c and $\boldsymbol{a}_{\varnothing}$ are both abstract, we must show $\forall F(cF \equiv \boldsymbol{a}_{\varnothing}F)$. Now, by theorem (266.3) and definition (263.1), we know $\neg\exists F\boldsymbol{a}_{\varnothing}F$. So by (103.9), we need only show $\neg\exists FcF$. For reductio, assume $\exists FcF$. Let P be an arbitrary such property, so that we know cP. But c is, by hypothesis, a bottom concept, and so $\forall d(c \preceq d)$. Since $\boldsymbol{a}_{\varnothing}$ is a concept, it follows that $c \preceq \boldsymbol{a}_{\varnothing}$. Hence, $\forall F(cF \to \boldsymbol{a}_{\varnothing}F)$. So $\boldsymbol{a}_{\varnothing}P$, given that cP, and thus $\exists F\boldsymbol{a}_{\varnothing}F$. Contradiction. ⋈

(652.3) (Exercise)

(652.4) Since the previous theorem is that $\boldsymbol{a}_\varnothing \prec \boldsymbol{a}_G$, the present theorem will follow by &I, ∃I, and the definition of the unique existence quantifier if we can establish that $\forall c(c \prec \boldsymbol{a}_G \rightarrow c = \boldsymbol{a}_\varnothing)$. So, by GEN, assume $c \prec \boldsymbol{a}_G$ and, for reductio, suppose $c \neq \boldsymbol{a}_\varnothing$. Since $\boldsymbol{a}_\varnothing$ encodes no properties, $c \neq \boldsymbol{a}_\varnothing$ implies $\exists F(cF \,\&\, \neg \boldsymbol{a}_\varnothing F)$. Suppose P is an arbitrary such property, so that we know cP and $\neg \boldsymbol{a}_\varnothing P$. Since $c \prec \boldsymbol{a}_G$ by assumption, it follows by definition of $\prec$ (649) both that:

(ϑ) $c \preceq \boldsymbol{a}_G$

(ξ) $c \neq \boldsymbol{a}_G$

From (ϑ) and cP, it follows that $\boldsymbol{a}_G P$, and so $P = G$, by the right conjunct of (426.1). Hence,

(ζ) cG

But (ϑ) and (ξ) jointly imply that there is a property that $\boldsymbol{a}_G$ encodes that c fails to encode. Suppose Q is such that (i) $\boldsymbol{a}_G Q$ and (ii) $\neg cQ$. Then from (i) and the right conjunct of (426.1), we know $Q = G$, and so from (ii), it follows that $\neg cG$, which contradicts (ζ). ⋈

(652.5) Assume *Bottom*(c). So by definition (651.1),

(ϑ) $\forall d(c \preceq d)$

To show *Atom*(c), we have to show $\neg\exists d(d \prec c)$, by (651.2). For reductio, assume $\exists d(d \prec c)$. Suppose d_1 is an arbitrary such concept, so that we know $d_1 \prec c$. Then by definition of $\prec$ (649), it follows that (a) $d_1 \preceq c$ and (b) $d_1 \neq c$. But if we instantiate d_1 into (ϑ), it follows that $c \preceq d_1$. But from (a) and this last fact, it follows by (626.1) that $d_1 = c$, contradicting (b). ⋈

(652.6) (Exercise)

(652.7) Given a proof of *Atom*($\boldsymbol{a}_\varnothing$) (652.6), it suffices, by &I, ∃I, and the definition of the unique existence quantifier, to show $\forall c(Atom(c) \rightarrow c = \boldsymbol{a}_\varnothing)$. By GEN, it suffices to show $Atom(c) \rightarrow c = \boldsymbol{a}_\varnothing$. So assume *Atom*($c$) and, for reductio, assume $c \neq \boldsymbol{a}_\varnothing$. From the atomicity of c, it follows that:

(ϑ) $\neg\exists d(d \prec c)$

Now $\boldsymbol{a}_\varnothing$ is a bottom concept (652.1), and so $\forall d(\boldsymbol{a}_\varnothing \preceq d)$, by definition (651.1). In particular, $\boldsymbol{a}_\varnothing \preceq c$. But from our reductio assumption, we also know $\boldsymbol{a}_\varnothing \neq c$, by symmetry of identity. Hence, $\boldsymbol{a}_\varnothing \prec c$, and so $\exists d(d \prec c)$, which contradicts (ϑ). ⋈

(654.1) Assume *Overlap*(c, d). Then by (653.1), $\exists F(cF \,\&\, dF)$. Assume P is such a property, so that we know $cP \,\&\, dP$. Now to show *Overlap**(c, d), we have to show $\exists e(e \preceq c \,\&\, e \preceq d)$. Let's show that Thin Form of P, namely $\boldsymbol{a}_P$, is a witness, so that it suffices to show both $\boldsymbol{a}_P \preceq c$ and $\boldsymbol{a}_P \preceq d$. To show that $\boldsymbol{a}_P \preceq c$, we have to show $\forall F(\boldsymbol{a}_P F \rightarrow cF)$, and by GEN, $\boldsymbol{a}_P F \rightarrow cF$. So assume $\boldsymbol{a}_P F$. Then by the second conjunct of (426.1), it follows that $F = P$. Since cP, it follows that cF. By analogous reasoning, $\boldsymbol{a}_P \preceq d$. ⋈

(654.2) Let's choose our witnesses to be, respectively, the null concept, $\boldsymbol{a}_\varnothing$, and the Thin Form of P, $\boldsymbol{a}_P$, where P is any arbitrarily chosen property. Then it suffices to show (a) $Overlap^*(\boldsymbol{a}_\varnothing, \boldsymbol{a}_P)$, and (b) $\neg Overlap(\boldsymbol{a}_\varnothing, \boldsymbol{a}_P)$. (a) By (625.1), we know $\boldsymbol{a}_\varnothing \preceq \boldsymbol{a}_\varnothing$, and by (652.3), we know $\boldsymbol{a}_\varnothing \preceq \boldsymbol{a}_P$. Hence, by &I and $\exists$I, $\exists e(e \preceq \boldsymbol{a}_\varnothing \,\&\, e \preceq \boldsymbol{a}_P)$, i.e., $Overlap^*(\boldsymbol{a}_\varnothing, \boldsymbol{a}_P)$. (b) By now familiar reasoning, $\neg\exists F \boldsymbol{a}_\varnothing F$. Hence, $\neg\exists F(\boldsymbol{a}_\varnothing F \,\&\, \boldsymbol{a}_P F)$. So $\neg Overlap(\boldsymbol{a}_\varnothing, \boldsymbol{a}_P)$. ⋈

(654.3) – **(654.7)** (Exercises)

(654.8) By GEN, it suffices to show $Overlap^*(c,d)$. Since $Bottom(\boldsymbol{a}_\varnothing)$ by (652.1), it follows by definition (651.1) that $\boldsymbol{a}_\varnothing$ is a part of every concept. Hence, we know both $\boldsymbol{a}_\varnothing \preceq c$ and $\boldsymbol{a}_\varnothing \preceq d$, for any concepts c and d. So, $\exists e(e \preceq c \,\&\, e \preceq d)$. Thus, $Overlap^*(c,d)$, by definition (653.2). ⋈

(654.9) We can show $\boldsymbol{a}_\varnothing$ is a witness by showing:

(a) $\boldsymbol{a}_\varnothing \preceq d$

(b) $\neg Overlap(\boldsymbol{a}_\varnothing, c)$

(a) follows immediately from the fact that $\boldsymbol{a}_\varnothing$ is a bottom concept, and so is a part of every concept, by (652.1) and (651.1). For (b), we know that $\neg\exists F \boldsymbol{a}_\varnothing F$. Hence, it follows *a fortiori* that $\neg\exists F(\boldsymbol{a}_\varnothing F \,\&\, cF)$. So $\neg Overlap(\boldsymbol{a}_\varnothing, c)$. ⋈

Alternatively, $d \ominus c$ is a witness.[475] Recall that $d \ominus c$ was defined in (646.5) (as part of Exercise 1) as $\imath e DifferenceOf(e,d,c)$, where $DifferenceOf(e,d,c)$ was defined in (646.1) as $\forall F(eF \equiv dF \,\&\, \neg cF)$. By &I and $\exists$I, it suffices to show:

(a) $d \ominus c \preceq d$

(b) $\neg Overlap(d \ominus c, c)$

For (a), we need to show $d \ominus cF \rightarrow dF$. So assume $d \ominus cF$. Then by Exercise (646.6), it follows that $dF \,\&\, \neg cF$. So dF. For (b), proceed by reductio. Assume $Overlap(d \ominus c, c)$. Then $\exists F(d \ominus cF \,\&\, cF)$. Suppose P is an arbitrary such property, so that we know $d \ominus cP$ and cP. From the former, it follows by Exercise (646.6) that $dP \,\&\, \neg cP$. So $\neg cP$. Contradiction. ⋈

(654.10) (Complete the proof sketched in the text.)

(655.1) – **(655.2)** (Exercises)

[475] Note that if $d \preceq c$, then $d \ominus c$ is $\boldsymbol{a}_\varnothing$, and so the proof reduces to the previous one.

(655.3) Assume $c \prec d$. Then, to show that $\boldsymbol{a}_\varnothing$ is a witness to $\exists e(e \prec d \,\&\, \neg Overlap(e,c))$, we show both (a) $\boldsymbol{a}_\varnothing \prec d$ and (b) $\neg Overlap(\boldsymbol{a}_\varnothing, c)$:

(a) By definition, we have to show both $\boldsymbol{a}_\varnothing \preceq d$ and $\boldsymbol{a}_\varnothing \neq d$. The former is a simple consequence of the fact that $\boldsymbol{a}_\varnothing$ is a bottom concept (652.1) and thus by definition (651.1) a part of every concept. For the latter, it suffices to show that d encodes a property, since we know, by theorem (266.3) and definition (263.1), that $\boldsymbol{a}_\varnothing$ doesn't. Note that from our initial hypothesis, it follows that $c \preceq d$ and $c \neq d$. But since the former implies $\forall F(cF \rightarrow dF)$ and the latter implies that one of c and d encodes a property the other fails to encode, it follows that $\exists F(dF \,\&\, \neg cF)$. *A fortiori*, $\exists F dF$.

(b) By now familiar reasoning, $\boldsymbol{a}_\varnothing$ encodes no properties. *A fortiori*, there is no property that both $\boldsymbol{a}_\varnothing$ encodes and c encodes. Hence $\neg Overlap(\boldsymbol{a}_\varnothing, c)$. ⋈

(657.1) – **(657.2)** (Exercises)

(660) By GEN, assume $Concept^+(x)$. Then by definition, $C!x$ and $\neg Null(x)$. From the latter, it follows that $\exists F xF$, by the definition of $(Null(x))$ in (263.1). And from the former, it follows that $\Box C!x$ (exercise). Now assume P is a witness to $\exists F xF$, so that we know xP. The $\Box xP$. Hence, $\exists F \Box xF$. So by the Buridan formula (168.1), $\Box \exists F xF$. Hence, we've established $\Box C!x \,\&\, \Box \exists F xF$. So, $\Box(C!x \,\&\, \exists F xF)$. Hence $\Box Concept^+(x)$. ⋈

(662.1) Clearly, since $\boldsymbol{a}_\varnothing$ fails to be a non-null concept, it fails to be a non-null bottom. ⋈

(662.2) Suppose, for reductio, that $\exists \underline{c} Bottom^+(\underline{c})$. Let $\underline{c}_1$ be such a non-null concept, so that we know $Bottom^+(\underline{c}_1)$, i.e., by (661), that $\forall \underline{d}(\underline{c}_1 \preceq \underline{d})$, i.e., that:

(ϑ) $\forall \underline{d} \forall F(\underline{c}_1 F \rightarrow \underline{d} F)$

Since $\underline{c}_1$ is non-null, $\exists F \underline{c}_1 F$. Suppose $\underline{c}_1 P$. Then consider $\boldsymbol{a}_{\overline{P}}$, i.e., the Thin Form of $\overline{P}$, which we know exists. Since $A!\boldsymbol{a}_{\overline{P}}$ (426.1), $C!\boldsymbol{a}_{\overline{P}}$ (612), and by (426.2), $\boldsymbol{a}_{\overline{P}}\overline{P}$. So $\boldsymbol{a}_{\overline{P}}$ is non-null concept. Hence (ϑ) implies $\forall F(\underline{c}_1 F \rightarrow \boldsymbol{a}_{\overline{P}} F)$. So since $\underline{c}_1 P$, it follows that $\boldsymbol{a}_{\overline{P}} P$. But $P \neq \overline{P}$ (199.5), and so it follows from the second conjunct of (426.1) that $\neg \boldsymbol{a}_{\overline{P}} P$. Contradiction. ⋈

(664.1) Clearly, $\boldsymbol{a}_G$ is a non-null concept. So, we have to show, by the definition of a non-null atom (663), that $\neg \exists \underline{d}(\underline{d} \prec \boldsymbol{a}_G)$. We do this with the help of a lemma, namely, if a non-null concept is a part of $\boldsymbol{a}_G$, it fails to be a proper part of $\boldsymbol{a}_G$:

(ϑ) $\forall \underline{d}(\underline{d} \preceq \boldsymbol{a}_G \rightarrow \neg \underline{d} \prec \boldsymbol{a}_G)$

> *Proof*. By GEN, we need to show $\underline{d} \preceq \boldsymbol{a}_G \to \neg \underline{d} \prec \boldsymbol{a}_G$. Assume $\underline{d} \preceq \boldsymbol{a}_G$. For reductio, assume $\underline{d} \prec \boldsymbol{a}_G$. Then by definition (649), we also know that $\underline{d} \neq \boldsymbol{a}_G$. So $\underline{d}$ and $\boldsymbol{a}_G$ must differ by one of their encoded properties. But since our first assumption implies $\forall F(\underline{d}F \to \boldsymbol{a}_G F)$, it must be that $\boldsymbol{a}_G$ encodes a property $\underline{d}$ doesn't encode, i.e., $\exists F(\boldsymbol{a}_G F \ \&\ \neg \underline{d}F)$. Suppose P is such a property, so that we know $\boldsymbol{a}_G P$ and $\neg \underline{d}P$. Since the former implies $P = G$ (426.1), it follows from the latter that $\neg \underline{d}G$. But $\underline{d}$ is, by hypothesis, non-null, and so $\exists F \underline{d}F$. Suppose Q is such a property, so that we know dQ. Then since $\underline{d} \preceq \boldsymbol{a}_G$, it follows that $\boldsymbol{a}_G Q$ and, by now familiar reasoning, $G = Q$. So it follows from the previously established fact $\neg \underline{d}G$ that $\neg \underline{d}Q$. Contradiction.

Now, for reductio, suppose that $\exists \underline{d}(\underline{d} \prec \boldsymbol{a}_G)$. Let $\underline{d}_1$ be such a non-null concept, so that we know $\underline{d}_1 \prec \boldsymbol{a}_G$. Then the definition of $\prec$ implies $\underline{d}_1 \preceq \boldsymbol{a}_G$. But then by ($\vartheta$), $\neg \underline{d}_1 \prec \boldsymbol{a}_G$. Contradiction. ⋈

(664.2) For reductio, assume $\exists! \underline{d}(\underline{d} \prec \underline{c})$. By definition of the unique existence quantifier, it follows that $\exists \underline{d}(\underline{d} \prec \underline{c} \ \&\ \forall \underline{e}(\underline{e} \prec \underline{c} \to \underline{e} = \underline{d}))$. Suppose $\underline{d}_1$ is such a non-null concept, so that we know:

(ϑ) $\underline{d}_1 \prec \underline{c} \ \&\ \forall \underline{e}(\underline{e} \prec \underline{c} \to \underline{e} = \underline{d}_1)$

By definition of $\prec$, the first conjunct of (ϑ) implies:

(ξ) $\underline{d}_1 \preceq \underline{c} \ \&\ \underline{d}_1 \neq \underline{c}$

By now familiar reasoning, (ξ) implies $\exists F(\underline{c}F \ \&\ \neg \underline{d}_1 F)$. Suppose P is such a property, so that we know both $\underline{c}P$ and $\neg \underline{d}_1 P$. Now consider the Thin Form of P, $\boldsymbol{a}_P$. Since $\boldsymbol{a}_P P$ (433.1) and $\neg \underline{d}_1 P$, it follows that:

(ζ) $\boldsymbol{a}_P \neq \underline{d}_1$

Now if we can establish that $\boldsymbol{a}_P \prec \underline{c}$, then we will have reached our contradiction, since this would imply, by the second conjunct of (ϑ), that $\boldsymbol{a}_P = \underline{d}_1$, contradicting ($\zeta$). So it remains show (a) $\boldsymbol{a}_P \preceq \underline{c}$ and (b) $\boldsymbol{a}_P \neq \underline{c}$:

(a) Assume $\boldsymbol{a}_P F$. Then by (426.1), $F = P$. But we already know $\underline{c}P$. So $\underline{c}F$.

(b) Assume, for reductio, $\boldsymbol{a}_P = \underline{c}$. From this and the first conjunct of (ϑ), it follows that $\underline{d}_1 \prec \boldsymbol{a}_P$. So $\exists \underline{d}(\underline{d} \prec \boldsymbol{a}_P)$. But by definition (663), this contradicts the atomicity$^+$ of $\boldsymbol{a}_P$ (664.1). ⋈

(664.3) Assume $Atom^+(\underline{c})$. Then, by definition:

(ϑ) $\neg \exists \underline{d}(\underline{d} \prec \underline{c})$

Assume $\underline{c}F$ & $\underline{c}G$ and, for reductio, $F \neq G$. Consider, then, the Thin Form of F, $\boldsymbol{a}_F$. Clearly, $\boldsymbol{a}_F$ is a non-null concept. Our strategy is to establish both that (a) $\boldsymbol{a}_F \preceq \underline{c}$ and (b) $\boldsymbol{a}_F \neq \underline{c}$, for then it follows by definition (649) that $\boldsymbol{a}_F \prec \underline{c}$ and thus $\exists \underline{d}(\underline{d} \prec \underline{c})$, contradicting ($\vartheta$):

(a) We have to show $\forall H(\boldsymbol{a}_F H \rightarrow \underline{c}H)$. So by GEN, assume $\boldsymbol{a}_F H$. Then by now familiar reasoning, it follows that $F = H$. Since we know $\underline{c}F$, it follows that $\underline{c}H$.

(b) By (426.1), we know $\boldsymbol{a}_F G \equiv G = F$. So it follows from our assumption that $F \neq G$ that $\neg \boldsymbol{a}_F G$. Since $\underline{c}G$ by assumption, it follows that $\boldsymbol{a}_F \neq \underline{c}$. ⋈

(664.4) (Exercise)

(665) ($\rightarrow$) Assume $Overlap(\underline{c}, \underline{d})$, i.e., that $\exists F(\underline{c}F \,\&\, \underline{d}F)$. Suppose P is such a property, so that we know both $\underline{c}P$ and $\underline{d}P$. Then choose our witness for e to be the Thin Form of P, i.e., $\boldsymbol{a}_P$, which is clearly a non-null concept. Since $\boldsymbol{a}_P$ encodes the single property P, it follows that both $\boldsymbol{a}_P \preceq \underline{c}$ and $\boldsymbol{a}_P \preceq \underline{d}$. Consequently, $\exists \underline{e}(\underline{e} \preceq \underline{c} \,\&\, \underline{e} \preceq \underline{d})$. ($\leftarrow$) Assume $\exists \underline{e}(\underline{e} \preceq \underline{c} \,\&\, \underline{e} \preceq \underline{d})$. Let $\underline{e}_1$ be such a concept, so that we know both $\underline{e}_1 \preceq \underline{c}$ and $\underline{e}_1 \preceq \underline{d}$. Since $\underline{e}_1$ is non-null, $\exists F \underline{e}_1 F$. Suppose P is such a property, so that $\underline{e}_1 P$. Then since both $\underline{e}_1 \preceq \underline{c}$ and $\underline{e}_1 \preceq \underline{d}$, it follows from each, respectively, that $\underline{c}P$ and $\underline{d}P$. So $\exists F(\underline{c}F \,\&\, \underline{d}F)$. ⋈

(667.1) We leave the proof of the first conjunct as an exercise. To prove the second conjunct, let P and Q be any two distinct properties and consider the Thin Forms of P and Q, $\boldsymbol{a}_P$ and $\boldsymbol{a}_Q$. Since both encode a property, both are non-null concepts. But there is no property that they both encode. Hence $\neg Overlap(\boldsymbol{a}_P, \boldsymbol{a}_Q)$. So $\exists \underline{c} \exists \underline{d} \neg Overlap(\underline{c}, \underline{d})$. ⋈

(667.2) – **(667.3)** (Exercises)

(667.4) Consider any three, pairwise distinct properties P, Q, and R. We know there are such by a previous theorem. Now consider the following three concepts: the Thin Form of P, i.e., $\boldsymbol{a}_P$; the (strictly canonical) concept that encodes just the two properties P and Q; and the (strictly canonical) concept that encodes just the two properties Q and R. Call the latter two concepts c_1 and c_2, respectively:

$$c_1 = \imath c \forall F(cF \equiv F = P \vee F = Q)$$

$$c_2 = \imath c \forall F(cF \equiv F = Q \vee F = R)$$

Clearly, $\boldsymbol{a}_P$, c_1, and c_2 are all non-null concepts. So it suffices by &I and $\exists$I to establish:

(a) $Overlap(\boldsymbol{a}_P, c_1)$

(b) $Overlap(c_1, c_2)$

(c) $\neg Overlap(\boldsymbol{a}_P, c_2)$

(a) Since $\boldsymbol{a}_P$ encodes just the property P and c_1 encodes just P and Q, there is a property they both encode, and so $Overlap(\boldsymbol{a}_P, c_1)$. (b) Since c_1 and c_2 both encode Q, $Overlap(c_1, c_2)$. (c) Since $\boldsymbol{a}_P$ encodes just P and c_2 encodes just Q and R, then there is no property they encode in common. Hence $\neg Overlap(\boldsymbol{a}_P, c_2)$. ⋈

(667.5) By classical quantificational reasoning, the claim we want to prove is equivalent to:

(ϑ) $\exists \underline{c} \exists \underline{d} \forall \underline{e}(\underline{e} \preceq \underline{d} \rightarrow Overlap(\underline{e}, \underline{c}))$

So we show, where G is any property, that $\boldsymbol{a}_G$ is a witness to both existential claims. Clearly, $\boldsymbol{a}_G$ is a non-null concept. So we have to show:

$$\forall \underline{e}(\underline{e} \preceq \boldsymbol{a}_G \rightarrow Overlap(\underline{e}, \boldsymbol{a}_G))$$

By GEN, it suffices to show $\underline{e} \preceq \boldsymbol{a}_G \rightarrow Overlap(\underline{e}, \boldsymbol{a}_G)$. So assume $\underline{e} \preceq \boldsymbol{a}_G$. Now to show $Overlap(\underline{e}, \boldsymbol{a}_G)$, we have to show $\exists F(\underline{e}F \,\&\, \boldsymbol{a}_G F)$. But G is a witness, as can be seen from the following reasoning. $\boldsymbol{a}_G G$ follows immediately (426.2). So it remains to show eG. Since $\underline{e}$ is non-null, $\exists F \underline{e}F$. Let H be such a property, so that we know eH. But since $\underline{e}$ is by hypothesis a part of $\boldsymbol{a}_G$, it follows that $\boldsymbol{a}_G H$. But then $H = G$ (426.1). So eG. ⋈

(667.6) (The proof was sketched in the text.)

(668.1) Assume $\underline{c} \prec \underline{d}$. Then $\underline{c} \preceq \underline{d} \,\&\, \underline{c} \neq \underline{d}$. So $\exists F(\underline{d}F \,\&\, \neg \underline{c}F)$. Let P be such a property, so that we know $\underline{d}P \,\&\, \neg \underline{c}P$. Then as our witness to $\exists \underline{e}(\underline{e} \preceq \underline{d} \,\&\, \neg Overlap(\underline{e}, \underline{c}))$, consider the Thin Form of P, $\boldsymbol{a}_P$, which is clearly a non-null concept. It suffices to show $\boldsymbol{a}_P \preceq \underline{d}$ and $\neg Overlap(\boldsymbol{a}_P, \underline{c})$. Since $\boldsymbol{a}_P$ encodes only P and d encodes P, $\underline{d}$ encodes every property $\boldsymbol{a}_P$ encodes. Hence, $\boldsymbol{a}_P \preceq \underline{d}$. Moreover, since $\boldsymbol{a}_P$ encodes just P and $\underline{c}$ doesn't encode P, there is no property that they encode in common. Hence $\neg Overlap(\boldsymbol{a}_P, \underline{c})$. ⋈

(668.2) Assume $\underline{d} \npreceq \underline{c}$. Then $\exists F(\underline{d}F \,\&\, \neg \underline{c}F)$ and our proof now reduces to that of the previous theorem. ⋈

(668.3) Assume $\underline{c} \prec \underline{d}$. Then $\underline{c} \preceq \underline{d} \,\&\, \underline{c} \neq \underline{d}$. So $\exists F(\underline{d}F \,\&\, \neg \underline{c}F)$. Let P be such a property, so that we know $\underline{d}P \,\&\, \neg \underline{c}P$. Now, in the proof of (668.1), we established that, under the same assumptions, the Thin Form of P, $\boldsymbol{a}_P$ is a witness to $\exists \underline{e}(\underline{e} \preceq \underline{d} \,\&\, \neg Overlap(\underline{e}, \underline{c}))$ by showing that both $\boldsymbol{a}_P \preceq \underline{d}$ and $\neg Overlap(\boldsymbol{a}_P, \underline{c})$. So to show that $\boldsymbol{a}_P$ is a witness to $\exists \underline{e}(\underline{e} \prec \underline{d} \,\&\, \neg Overlap(\underline{e}, \underline{c}))$, it suffices to show only $\boldsymbol{a}_P \neq \underline{d}$.[476] For reductio, suppose $\boldsymbol{a}_P = \underline{d}$. Then from our initial assumption,

[476]**Exercise**. Under the assumption that $c \prec d$, show that $\underline{d} \ominus \underline{c}$ is also a witness to $\exists \underline{e}(\underline{e} \prec \underline{d} \,\&\, \neg Overlap(\underline{e}, \underline{c}))$.

it follows that $\underline{c} \prec \boldsymbol{a}_P$. Hence $\exists \underline{e}(\underline{e} \prec \boldsymbol{a}_P)$. But by theorem (664.1), we know that $Atom^+(\boldsymbol{a}_P)$. Hence, by the definition of a non-null atom (663), it follows that $\neg\exists \underline{e}(\underline{e} \prec \boldsymbol{a}_P)$. Contradiction. ⋈

(671.1) – (671.3) (Exercises)

(674) (Exercise)

(675.1) – (675.2) (Exercises)

(676) We want to show $\boldsymbol{c}_G \oplus \boldsymbol{c}_H = c_1$, where c_1 is the following canonical concept:

$$c_1 = \imath c\forall F(cF \equiv G \Rightarrow F \vee H \Rightarrow F)$$

Let φ be the condition $G \Rightarrow F \vee H \Rightarrow F$. Then we leave it as a simple exercise to show that φ is a rigid condition on properties and, hence, that c_1 is strictly canonical. From this, it follows by theorem (261.2) that:

$(\vartheta)\ \ \forall F(c_1F \equiv G \Rightarrow F \vee H \Rightarrow F)$

Now to see $\boldsymbol{c}_G \oplus \boldsymbol{c}_H$ and c_1 encode the same properties, we establish $\boldsymbol{c}_G \oplus \boldsymbol{c}_H F \equiv c_1 F$:

$$\begin{array}{rcll}
\boldsymbol{c}_G \oplus \boldsymbol{c}_H F & \equiv & \boldsymbol{c}_G F \vee \boldsymbol{c}_H F & \text{by (620.3)} \\
 & \equiv & G \Rightarrow F \vee \boldsymbol{c}_H F & \text{by (675.1), (88.8.h)} \\
 & \equiv & G \Rightarrow F \vee H \Rightarrow F & \text{by (675.1), (88.8.g)} \\
 & \equiv & c_1 F & \text{by } (\vartheta)
\end{array}$$

⋈

(677.1) – (677.2) (Exercises)

(673) (Exercise)

(681.1) – (681.3) (Exercises)

(683) Our theorem is implied by theorem (205.1), i.e., $\Diamond\exists x(E!x \,\&\, \Diamond\neg E!x)$, as follows:

$$\begin{array}{rcll}
\Diamond\exists x(E!x \,\&\, \Diamond\neg E!x) & \rightarrow & \exists x\Diamond(E!x \,\&\, \Diamond\neg E!x) & \text{by BF}\Diamond \\
 & \rightarrow & \exists x(\Diamond E!x \,\&\, \Diamond\neg E!x) & \text{by (165.11) and (160.2)} \\
 & \rightarrow & \exists u(\Diamond E!u \,\&\, \Diamond\neg E!u) & \text{(exercise)} \\
 & \rightarrow & \exists u\Diamond(E!u \,\&\, \Diamond\neg E!u) & \text{by (165.11) and (160.2)} \\
 & \rightarrow & \exists F\exists u\Diamond(Fu \,\&\, \Diamond\neg Fu) & \text{by } \exists\text{I} \\
 & \rightarrow & \exists u\exists F\Diamond(Fu \,\&\, \Diamond\neg Fu) & \text{(exercise)}
\end{array}$$

⋈

(685)★ By (681.3), $\imath cConceptOf(c,u)\downarrow$. So by our theory of definitions, $\boldsymbol{c}_u = \imath cConceptOf(c,u)$. Then by (145.2)★, it follows that $ConceptOf(\boldsymbol{c}_u,u)$. Hence, by definition (680), $\forall F(\boldsymbol{c}_u F \equiv Fu)$. (See foonote 341.) Instantiating to G yields $\boldsymbol{c}_u G \equiv Gu$. ⋈

(686) We prove this theorem with the help of the following Lemma, the proof of which was the subject of the Exercise in (685):

> *Lemma*: $\forall F(c_u F \equiv \mathcal{A}Fu)$
> *Proof*: If we eliminate the restricted variable, then we want to show: $O!y \to \forall F(c_y F \equiv \mathcal{A}Fy)$. So assume $O!y$. Since F isn't free in our assumption, it suffices to show $c_y F \equiv \mathcal{A}Fy$, by GEN. Now, as an instance of (258), we know $\imath x(A!x \& \forall F(xF \equiv Fy))F \equiv \mathcal{A}Fy$. But if we eliminate the restricted variable in the fact cited at the outset of Remark (684), then our assumption $O!y$ implies $c_y = \imath x(A!x \& \forall F(xF \equiv Fy))$. So by Rule =E, $c_y F \equiv \mathcal{A}Fy$.

This Lemma is needed only for the left-to-right direction of our theorem.

($\to$) Assume $\boldsymbol{c}_u G$. By definition of $\succeq$, we need to show $\boldsymbol{c}_G \preceq \boldsymbol{c}_u$, and by definition of $\preceq$, show $\forall F(\boldsymbol{c}_G F \to \boldsymbol{c}_u F)$. Since F isn't free in our assumption, it suffices to show $\boldsymbol{c}_G F \to \boldsymbol{c}_u F$. So assume $\boldsymbol{c}_G F$. Then by (675.1), we know $G \Rightarrow F$, i.e., that $\Box\forall x(Gx \to Fx)$. By theorem (132), it follows that $\mathcal{A}\forall x(Gx \to Fx)$, and by axiom (44.3), that $\forall x\mathcal{A}(Gx \to Fx)$. Thus $\mathcal{A}(Gu \to Fu)$, and so $\mathcal{A}Gu \to \mathcal{A}Fu$, by (131). But from our initial assumption that $\boldsymbol{c}_u G$, it follows by our Lemma that $\mathcal{A}Gu$. Hence, $\mathcal{A}Fu$. But, then again by our Lemma, $c_u F$.

($\leftarrow$) Assume $\boldsymbol{c}_u \succeq \boldsymbol{c}_G$, i.e., $\boldsymbol{c}_G \preceq \boldsymbol{c}_u$, i.e., $\forall F(\boldsymbol{c}_G F \to \boldsymbol{c}_u F)$. Now independently, by (675.2), we know $\boldsymbol{c}_G G$. Hence, $\boldsymbol{c}_u G$. ⋈

(688) By definition (687) we have to show $\forall F(cF \vee c\overline{F})$, and so by GEN, it suffices to show $\boldsymbol{c}_u F \vee \boldsymbol{c}_u \overline{F}$. We establish this by disjunctive syllogism (86.3.c) from $\mathcal{A}Fu \vee \mathcal{A}\neg Fu$, which is an instance of theorem (139.1). We reason from each disjunct using the Lemma established at the outset of the proof of (686), namely, $\forall F(c_u F \equiv \mathcal{A}Fu)$. If $\mathcal{A}Fu$, then $c_u F$, by our Lemma. If $\mathcal{A}\neg Fu$, then note that since $\overline{F}x \equiv \neg Fx$ is, by theorem (199.1), a modally strict theorem that holds universally for every object, it holds for ordinary objects, so that $\overline{F}u \equiv \neg Fu$ is a modally strict theorem. So we can use its commuted form with a Rule of Substitution to conclude $\mathcal{A}\overline{F}u$. ⋈

(690.1)★ – **(690.2)**★ (Exercise)

(691.1) ($\to$) Assume $\boldsymbol{c}_{\forall G}F$. Independently, by now familiar reasoning, it follows from an appropriate instance of (258) and definition (689), by Rule =E and GEN, that:

(ϑ) $\forall F(c_{\forall G}F \equiv \mathcal{A}\forall x(Gx \to Fx))$

So our assumption implies, by (ϑ), that $\mathcal{A}\forall x(Gx \to Fx)$. Now by the definitions of $\succeq$ and $\preceq$, we have to show $\forall H(\boldsymbol{c}_F H \to \boldsymbol{c}_{\forall G}H)$. So, by GEN, assume $\boldsymbol{c}_F H$, to show $\boldsymbol{c}_{\forall G}H$. But from $\boldsymbol{c}_F H$, we know by (675.1) that $F \Rightarrow H$, from which it follows by definition of $\Rightarrow$ and (132) that $\mathcal{A}\forall x(Fx \to Hx)$. We leave it as an exercise to show that the two facts we've established, namely $\mathcal{A}\forall x(Gx \to Fx)$ and $\mathcal{A}\forall x(Fx \to Hx)$, imply $\mathcal{A}\forall x(Gx \to Hx)$. So by ($\vartheta$), $\boldsymbol{c}_{\forall G}H$. ($\leftarrow$) Assume $\boldsymbol{c}_{\forall G} \succeq \boldsymbol{c}_F$. By the definitions of $\succeq$ and $\preceq$, this implies $\forall H(\boldsymbol{c}_F H \to \boldsymbol{c}_{\forall G}H)$. As an instance of this latter, we know $\boldsymbol{c}_F F \to \boldsymbol{c}_{\forall G} F$, from which it follows, by (675.2), that $\boldsymbol{c}_{\forall G} F$. ⋈

(691.2) ($\to$) Assume $\boldsymbol{c}_{\exists G} F$. Independently, by now familiar reasoning, it follows from an appropriate instance of (258) and definition (689), by Rule =E and GEN, that:

(ϑ) $\forall F(c_{\exists G} F \equiv \mathcal{A}\exists x(Gx \,\&\, Fx))$

So our assumption implies, by (ϑ), that $\mathcal{A}\exists x(Gx \,\&\, Fx)$. Now by the definitions of $\succeq$ and $\preceq$, we have to show $\forall H(\boldsymbol{c}_F H \to \boldsymbol{c}_{\exists G} H)$. So, by GEN, assume $\boldsymbol{c}_F H$, to show $\boldsymbol{c}_{\exists G} H$. But from $\boldsymbol{c}_F H$, we know by (675.1) that $F \Rightarrow H$, from which it follows by definition of $\Rightarrow$ and (132) that $\mathcal{A}\forall x(Fx \to Hx)$. We leave it as an exercise to show that the two facts we've established, namely $\mathcal{A}\exists x(Gx \& Fx)$ and $\mathcal{A}\forall x(Fx \to Hx)$, imply $\mathcal{A}\exists x(Gx \,\&\, Hx)$. So by ($\vartheta$), $\boldsymbol{c}_{\exists G} H$. ($\leftarrow$) Assume $\boldsymbol{c}_{\exists G} \succeq \boldsymbol{c}_F$. By the definitions of $\succeq$ and $\preceq$, our assumption implies $\forall H(\boldsymbol{c}_F H \to \boldsymbol{c}_{\exists G} H)$. As an instance of this latter, we know $\boldsymbol{c}_F F \to \boldsymbol{c}_{\exists G} F$, from which it follows, by (675.2), that $\boldsymbol{c}_{\exists G} F$. ⋈

(692.1)★ By the commutativity of the biconditional, it follows from (685)★ that $Gu \equiv \boldsymbol{c}_u G$. But from this and (686), and it follows that $Gu \equiv \boldsymbol{c}_u \succeq \boldsymbol{c}_G$ by a biconditional syllogism. ⋈

(692.2)★ – **(692.3)**★ (Exercises)

(699.1) Assume the antecedent and let a and w_1 be such an ordinary object and possible world, so that we know $RealizesAt(a, c, w_1)$ and $RealizesAt(a, d, w_1)$. Then, by the definition of realization (697), we know both $\forall F(w_1 \models Fa \equiv cF)$ and $\forall F(w_1 \models Fa \equiv dF)$. So, by the laws of quantified biconditionals, it follows that $\forall F(cF \equiv dF)$. Since c and d are concepts, they are abstract (612). So, by theorem (245.2), $c = d$. ⋈

(699.2) Assume the antecedent and let c_1 and w_1 be such a concept and possible world, so that we know $RealizesAt(u, c_1, w_1)$ and $RealizesAt(v, c_1, w_1)$. Then, by the definition of realization (697), we know both $\forall F(w_1 \models Fu \equiv c_1 F)$ and $\forall F(w_1 \models Fv \equiv c_1 F)$. So by the laws of quantified biconditionals, we know: $\forall F(w_1 \models Fu \equiv w_1 \models Fv)$. Instantiating to F, it follows that $(w_1 \models Fu) \equiv (w_1 \models Fv)$. But then by the right-to-left direction of (545.4), $w_1 \models (Fu \equiv Fv)$. Since F is arbitrary, it follows by Rule $\forall$I that $\forall F(w_1 \models (Fu \equiv Fv))$. So by the right-to-left direction of (545.5), $w_1 \models \forall F(Fu \equiv Fv)$. Since u and v are, by hypothesis, ordinary objects, they are necessarily such (180.1), i.e., $\Box O!u$ and $\Box O!v$, and so by a fundamental theorem of world theory (543.2), $\forall w(w \models O!u)$ and $\forall w(w \models O!v)$. In particular, $w_1 \models O!u$ and $w_1 \models O!v$. But we only need the first of these facts, since we may infer from it the following, in sequence, by the laws governing possible worlds ((545.3) and (545.1):

$$w_1 \models (O!u \lor O!v)$$

$$w_1 \models ((O!u \lor O!v) \,\&\, \forall F(Fu \equiv Fv))$$

Now independently, export the antecedent of the consequent of theorem (242.1) and apply RN, and we obtain, as a modally strict theorem:

$$\Box[\big((O!x \lor O!y) \,\&\, \forall F(Fx \equiv Fy)\big) \rightarrow x =_E y]$$

Since worlds are 1-modally closed (528), it follows from our last two displayed results that $w_1 \models u =_E v$. Hence, $\exists w(w \models u =_E v)$, and so $\Diamond u =_E v$, by (543.1). It follows that $u =_E v$ (234.2), and then $u = v$ (233.2). ⋈

(699.3) Assume the antecedent and suppose that a and c_1 are such an ordinary object and concept, so that we know $RealizesAt(a, c_1, w)$ and $RealizesAt(a, c_1, w')$. So we know, by the definition of realization (697), that $\forall F(w \models Fa \equiv c_1 F)$ and $\forall F(w' \models Fa \equiv c_1 F)$. So, by the laws of quantified biconditionals, we know:

(ϑ) $\forall F(w \models Fa \equiv w' \models Fa)$

Suppose, for reductio, that $w \neq w'$. Then, since w and w' are possible worlds, and hence situations, we know by (474) that there must be a proposition, say q_1, true at one and not at the other. Without loss of generality, assume $w \models q_1$ and $w' \not\models q_1$. From the former, it follows by a useful fact about possible worlds (551), that $w \models [\lambda y\, q_1]a$. So, by ($\vartheta$), $w' \models [\lambda y\, q_1]a$. So again by (551), $w' \models q_1$. Contradiction. ⋈

(701) Assume $AppearsAt(c, w)$. By the definition of appearance (700), it follows that some ordinary individual, say a, is such that $RealizesAt(a, c, w)$. It therefore remains to show uniqueness, i.e., that $\forall v(RealizesAt(v, c, w) \rightarrow v = a)$. So, by GEN, assume $RealizesAt(v, c, w)$. Then the antecedent of (699.2) holds and we can conclude $v = a$. ⋈

(702) Assume $AppearsAt(c, w)$. By the definition of appearance (700), it follows that some ordinary individual, say a, is such that $RealizesAt(a, c, w)$. So, by definition (697):

(ϑ) $\forall F(w \models Fa \equiv cF)$

Since we know $O!a$, we therefore know $\Box O!a$, by (180.1). Hence by a fundamental theorem of world theory (543.2), $\forall w'(w' \models O!a)$. In particular, $w \models O!a$. Hence, by (ϑ), $cO!$. ⋈

(704) Suppose $AppearsAt(c, w)$. So some ordinary object, say a, realizes c at w, i.e.,

(ϑ) $\forall F(w \models Fa \equiv cF)$

By (703) and GEN, we have to show that $c\Sigma p \equiv w \models p$. ($\rightarrow$) Assume $c\Sigma p$. Then by (295), $c[\lambda y\, p]$. So by (ϑ), $w \models [\lambda y\, p]a$. Hence, by (551), $w \models p$. ($\leftarrow$) Assume $w \models p$. Then by (551), $w \models [\lambda y\, p]a$. Hence by ($\vartheta$), $c[\lambda y\, p]$ and (295) now implies $c\Sigma p$. ⋈

(705) Assume the antecedent and suppose c_1 is such a concept, so that we know $AppearsAt(c_1, w)$ and $AppearsAt(c_1, w')$. Then by (704), it follows, respectively, that $Mirrors(c_1, w)$ and $Mirrors(c_1, w')$. We may infer from these, respectively, by the definition of mirroring (703), that $\forall p(c_1\Sigma p \equiv w \models p)$ and $\forall p(c_1\Sigma p \equiv w' \models p)$. So by the laws of quantified biconditionals, we know $\forall p(w \models p \equiv w' \models p)$. But since w and w' are both possible worlds, and hence situations, it follows by a fact about the identity of situations (474), that $w = w'$.

(706) We prove only the left-to-right direction, since the right-to-left direction is an instance of the T schema. To simplify the proof of the left-to-right direction, we first establish the following lemma:

(ϑ) $(w \models Fu \equiv cF) \equiv \Box(w \models Fu \equiv cF)$

> *Proof*. By axiom (51), we know both $w[\lambda y\, Fu] \rightarrow \Box w[\lambda y\, Fu]$ and $cF \rightarrow \Box cF$ are modally strict theorems. Hence by expanded RN (341.3.a) and &I, it follows that:
>
> $$\Box(w[\lambda y\, Fu] \rightarrow \Box w[\lambda y\, Fu])\ \&\ \Box(cF \rightarrow \Box cF)$$
>
> Hence by theorem (172.5) and the T schema (45.2), it follows that:
>
> $$(w[\lambda y\, Fu] \equiv cF) \rightarrow \Box(w[\lambda y\, Fu] \equiv cF)$$
>
> But since the converse of the above holds by the T schema, we may infer:
>
> (ξ) $(w[\lambda y\, Fu] \equiv cF) \equiv \Box(w[\lambda y\, Fu] \equiv cF)$
>
> Now given the discussion in Remark (515), it is a modally strict fact that $w[\lambda y\, Fu] \equiv w \models Fu$. So ($\vartheta$) follows from ($\xi$) by a Rule of Substitution.

Since (ϑ) is a modally strict theorem, a Rule of Substitution allows one to substitute instances of the right condition for the corresponding instances of the left condition whenever the left condition occurs as a subformula, and vice versa. Hence we may reason as follows:

$$\begin{array}{llll}
AppearsAt(c,w) & \rightarrow & \exists u RealizesAt(u,c,w) & \text{By df (700)} \\
& \rightarrow & RealizesAt(a,c,w) & \text{`}a\text{' arbitrary and ordinary} \\
& \rightarrow & \forall F(w \models Fa \equiv cF) & \text{by df (697)} \\
& \rightarrow & \forall F \Box (w \models Fa \equiv cF) & \text{by } (\vartheta)\text{, Rule of Substitution} \\
& \rightarrow & \Box \forall F(w \models Fa \equiv cF) & \text{by BF (167.1)} \\
& \rightarrow & \Box RealizesAt(a,c,w) & \text{by df (697)} \\
& \rightarrow & \exists u \Box RealizesAt(u,c,w) & \text{by } \exists\text{I and } \exists\text{E} \\
& \rightarrow & \Box \exists u RealizesAt(u,c,w) & \text{by Buridan (168.1)} \\
& \rightarrow & \Box AppearsAt(c,w) & \text{by df (700)}
\end{array}$$

(707) Assume the antecedent and suppose c_1 is such a concept, so that we know both $RealizesAt(u,c_1,w)$ and $RealizesAt(v,c_1,w')$. By respective applications of $\exists$I, it follows that $AppearsAt(c_1,w)$ and $AppearsAt(c_1,w')$. Hence $w = w'$, by (705). From this and $RealizesAt(v,c_1,w')$, we know $RealizesAt(v,c_1,w)$. From this and $RealizesAt(u,c_1,w)$, it follows, by (699.2), that $u = v$. ⋈

(708.1) As an instance of (538) we know $\boldsymbol{w}_\alpha \models Fu \equiv \mathcal{A}Fu$. The lemma established at the beginning of the proof of (686) implies $c_u F \equiv \mathcal{A}Fu$, i.e., $\mathcal{A}Fu \equiv c_u F$. So $\boldsymbol{w}_\alpha \models Fu \equiv c_u F$. By GEN, $\forall F(\boldsymbol{w}_\alpha \models Fu \equiv c_u F)$. So by definition of realization (697), $Realizes(u, \boldsymbol{c}_u, \boldsymbol{w}_\alpha)$. ⋈

(708.2) – **(708.3)** (Exercises)

(710) Assume $\exists u ConceptOf(c,u)$. Suppose a is such an ordinary object, so that we know $ConceptOf(c,a)$. Then by definition (680), it follows that:

(ϑ) $\forall F(cF \equiv Fa)$

By (709), we have to show $\exists w AppearsAt(c,w)$. So by (700), we have to show $\exists w \exists u RealizesAt(u,c,w)$, and so by (697), we have to show: $\exists w \exists u \forall F(w \models Fu \equiv cF)$. Thus, if we can find witnesses to both existential quantifiers in this last claim, we're done.

By the simplified comprehension principle for situations (486.1), there is a situation that makes true encodes all and only the propositions p such that for some property F that a exemplifies, F is the property $[\lambda y\, p]$:

$$\exists s \forall p(s \models p \equiv \exists F(Fa \,\&\, F = [\lambda y\, p]))$$

Let s_1 be such a situation, so that we know:

(ξ) $\forall p(s_1 \models p \equiv \exists F(Fa \,\&\, F = [\lambda y\, p]))$

If we can show (a) s_1 is a possible world, and (b) $\forall F(s_1 \models Fa \equiv cF)$, then s_1 and a are the two witnesses we need.

(a) Since s_1 is a situation, all we have to do to show s_1 is a possible world is to show $\Diamond \forall q(s_1 \models q \equiv q)$. By the T$\Diamond$ schema, it suffices to show $\forall q(s_1 \models q \equiv q)$.

By GEN, it suffices to show $s_1 \models q \equiv q$. ($\rightarrow$) Assume $s_1 \models q$. Hence by (ξ), $\exists F(Fa \,\&\, F = [\lambda y\, q])$. Suppose P is such a property, so that we know $Pa \,\&\, P = [\lambda y\, q]$. Then $[\lambda y\, q]a$, by substitution of identicals. So by β-Conversion, q. ($\leftarrow$) Assume q. Then $[\lambda y\, q]a$, by β-Conversion. So by =I, we may conclude $[\lambda y\, q]a \,\&\, [\lambda y\, q] = [\lambda y\, q]$. Hence, $\exists F(Fa \,\&\, F = [\lambda y\, q])$. So by ($\xi$), $s_1 \models q$.

(b) By GEN, we have to show $s_1 \models Fa \equiv cF$. To do this, we begin by noting that as part of (a), we established $\forall q(s_1 \models q \equiv q)$. If we instantiate this to the term Fa, it follows that $s_1 \models Fa \equiv Fa$. So by GEN, $\forall F(s_1 \models Fa \equiv Fa)$. But ($\vartheta$) is equivalent to $\forall F(Fa \equiv cF)$. So by the laws of quantified biconditionals, $\forall F(s_1 \models Fa \equiv cF)$. $\bowtie$

(711) By (708.2), $AppearsAt(\boldsymbol{c}_u, \boldsymbol{w}_\alpha)$. So $\exists w(AppearsAt(\boldsymbol{c}_u, w))$. By definition $PossibleIndividualConcept(\boldsymbol{c}_u)$. $\bowtie$

(712) By GEN, assume $PossibleIndividualConcept(x)$. Then, if we eliminate the restricted variable in definition (709), we know:

(ϑ) $C!x \,\&\, \exists y(PossibleWorld(y) \,\&\, AppearsAt(x, y))$

We want to show $\Box PossibleIndividualConcept(x)$, i.e.,

$$\Box(C!x \,\&\, \exists y(PossibleWorld(y) \,\&\, AppearsAt(x, y)))$$

So by (158.3), it suffices to show:

$$\Box C!x \,\&\, \Box\exists y(PossibleWorld(y) \,\&\, AppearsAt(x, y)))$$

But $\Box C!x$ follows from the first conjunct of (ϑ) (exercise). So by the Buridan formula (168.1), it remains and suffices to show:

(ξ) $\exists y \Box(PossibleWorld(y) \,\&\, AppearsAt(x, y))$

Now let w be a witness to the second conjunct of (ϑ), so that we know both $PossibleWorld(w)$ and $AppearsAt(x, w)$. Then we establish (ξ) by showing w is a witness. So by (158.3), we have to show:

$$\Box PossibleWorld(w)$$

$$\Box AppearsAt(x, w)$$

The first is easy: $\Box PossibleWorld(w)$ follows from $PossibleWorld(w)$ by (513). To show that $\Box AppearsAt(x, w)$ follows from $AppearsAt(x, a)$, note that $AppearsAt$ is defined in terms of restricted variables and so when we expand definitions (700) and (697), we have:

$$AppearsAt(x, w) \equiv_{df} C!x \,\&\, PossibleWorld(w) \,\&\, \exists z(O!z \,\&\, RealizesAt(z, x, w))$$

$$RealizesAt(z, x, w) \equiv_{df} O!z \,\&\, C!x \,\&\, PossibleWorld(w) \,\&\, \forall F((w \models Fz) \equiv xF)$$

Hence, $AppearsAt(x,w)$ implies, once we've expanded it by definition and simplified:

(ζ) $C!x \mathbin{\&} PossibleWorld(w) \mathbin{\&} \exists z(O!z \mathbin{\&} \forall F((w \models Fz) \equiv xF))$

Since we know the first two conjuncts are necessary, to prove that the entire claim is necessary, it remains only to show the following, by (158.3):

$$\Box\exists z(O!z \mathbin{\&} \forall F((w \models Fz) \equiv xF))$$

Again, by the Buridan formula (168.1), it suffices to show:

(ω) $\exists z\Box(O!z \mathbin{\&} \forall F((w \models Fz) \equiv xF))$

Now let b be a witness to the third conjunct of (ζ), so that we know $O!b$ and $\forall F((w \models Fb) \equiv xF)$. So to see that b is the needed witness to (ω), it suffices, again by (158.3), to show $\Box O!b$ and $\Box\forall F((w \models Fb) \equiv xF)$. But $\Box O!b$ follows from $O!b$ (180.1). To show the latter, it suffices, by the Barcan Formula (167.1), to show $\forall F\Box((w \models Fb) \equiv xF)$. But from $\forall F((w \models Fb) \equiv xF)$ it follows that $(w \models Fb) \equiv xF$. Since w is a situation, the left condition implies, by (471), that $w[\lambda z\, Fb]$. So $w[\lambda z\, Fb] \equiv xF$. Hence by (179.6), $\Box(w[\lambda z\, Fb] \equiv xF)$. Since F isn't free in any assumption, it follows that $\forall F\Box(w[\lambda z\, Fb] \equiv xF)$. ⋈

(713.1) Consider any possible-individual concept $\hat{c}$. Then by definition (709), there is some possible world, say w_1, such that $AppearsAt(\hat{c},w_1)$. It therefore remains to show that $\forall w(AppearsAt(\hat{c},w) \to w = w_1)$, and by GEN, that $AppearsAt(\hat{c},w) \to w = w_1$. So assume that $AppearsAt(\hat{c},w)$. Then by a fact about appearance (705), it follows that $w = w_1$. ⋈

(713.2) By theorem (706), we know:

$$AppearsAt(c,w) \equiv \Box AppearsAt(c,w)$$

Since this holds for arbitrary concepts, it holds for possible-individual concepts:

$$AppearsAt(\hat{c},w) \equiv \Box AppearsAt(\hat{c},w)$$

Since this is a modally strict equivalence, it follows from the previous theorem (713.1) by a Rule of Substitution that:

$$\exists! w\Box AppearsAt(\hat{c},w) \qquad ⋈$$

[Alternatively, $AppearsAt(c,w) \to \Box AppearsAt(c,w)$ follows *a fortiori* from theorem (706). *A fortiori*, $AppearsAt(\hat{c},w) \to \Box AppearsAt(\hat{c},w)$. So by GEN, it follows that $\forall w(AppearsAt(\hat{c},w) \to \Box AppearsAt(\hat{c},w))$. From this result and the previous theorem (713.1), it follows by theorem (129) that $\exists! w\Box AppearsAt(\hat{c},w)$.]

(713.3) (Exercise)

(715.1) As an instance of (153.1), we know:

$$\exists! w \Box AppearsAt(\hat{c}, w) \rightarrow \forall y(y = \imath w(AppearsAt(\hat{c}, w)) \rightarrow AppearsAt(\hat{c}, y))$$

So by (713.2), it follows that:

$$\forall y(y = \imath w(AppearsAt(\hat{c}, w)) \rightarrow AppearsAt(\hat{c}, y))$$

We may instantiate this to $\boldsymbol{w}_{\hat{c}}$ since we know it exists, to obtain:

$$\boldsymbol{w}_{\hat{c}} = \imath w(AppearsAt(\hat{c}, w) \rightarrow AppearsAt(\hat{c}, \boldsymbol{w}_{\hat{c}})$$

But then by definition (714), it follows that $AppearsAt(\hat{c}, \boldsymbol{w}_{\hat{c}})$. ⋈

(715.2) By (715.1) and (704). ⋈

(715.3) By the definitions of $\geq$ (624.2) and $\preceq$ (624.1), and GEN, we have to show: $\boldsymbol{w}_{\hat{c}}F \rightarrow \hat{c}F$. So assume $\boldsymbol{w}_{\hat{c}}F$. Since $\boldsymbol{w}_{\hat{c}}$ is a situation, every property that it encodes is a propositional property, and so for some proposition, say q_1, $F = [\lambda y\, q_1]$. Hence, $\boldsymbol{w}_{\hat{c}}[\lambda y\, q_1]$. By definition of $\models$, this implies that $\boldsymbol{w}_{\hat{c}} \models q_1$. Now, independently, we know by (715.2) that $Mirrors(\hat{c}, \boldsymbol{w}_{\hat{c}})$. So by definition of mirroring, $\forall p(\hat{c}\Sigma p \equiv \boldsymbol{w}_{\hat{c}} \models p)$. Since we've established $\boldsymbol{w}_{\hat{c}} \models q_1$, it follows that $\hat{c}\Sigma q_1$. By definition of Σ, this implies $\hat{c}[\lambda y\, q_1]$, i.e., $\hat{c}F$. ⋈

(716) ($\rightarrow$) Assume $\hat{c}G$. Before we show $\hat{c} \geq \boldsymbol{c}_G$, note that since $\hat{c}$ is a possible-individual concept, it follows by definition (709) that for some possible world, say w_1, $AppearsAt(\hat{c}, w_1)$. Further, by the definition of appearance (700), for some ordinary object, say a, we know $RealizesAt(a, \hat{c}, w_1)$. So by definition (697), this implies:

(ϑ) $\forall H(w_1 \models Ha \equiv \hat{c}H)$

Now by the definition of $\geq$ (624.2) and $\preceq$ (624.1), to show $\hat{c} \geq \boldsymbol{c}_G$, we have to show $\forall F(\boldsymbol{c}_G F \rightarrow \hat{c}F)$. So, by GEN, assume $\boldsymbol{c}_G F$. It follows that $\Box\forall x(Gx \rightarrow Fx)$, by (675.1) and (442). So by a Fundamental Theorem of Possible World Theory (543.2), we have $\forall w(w \models \forall x(Gx \rightarrow Fx))$. In particular, $w_1 \models \forall x(Gx \rightarrow Fx)$. But since $\hat{c}G$ is our global assumption, it follows from (ϑ) that $w_1 \models Ga$. Since the conjunction of $\forall x(Gx \rightarrow Fx)$ and Ga necessarily implies Fa, and possible worlds are binary-closed under necessary implication (528), it follows that $w_1 \models Fa$. Hence, by (ϑ), it follows that $\hat{c}F$.

($\leftarrow$) Assume $\hat{c} \geq \boldsymbol{c}_G$. Then, by now well-known reasoning, $\forall F(\boldsymbol{c}_G F \rightarrow \hat{c}F)$. But we also know $\boldsymbol{c}_G G$ (675.2). So $\hat{c}G$. ⋈

(717.1) Before we begin the proof proper, we establish two preliminary facts. This first fact concerns arbitrary witnesses to existential quantifiers. If $\hat{c}$ is a possible-individual concept, then by definition of (709), we $\exists w AppearsAt(\hat{c}, w)$. Suppose w_1 is such a world, so that we know $AppearsAt(\hat{c}, w_1)$. Then, by definition of appears at (700), $\exists u RealizesAt(u, \hat{c}, w_1)$, and by definition of realizes at (697), $\exists u \forall F(w_1 \models Fu \equiv \hat{c}F)$. Suppose a is such an ordinary object. So it is a fact about arbitrary witnesses w_1 and a that:

(ϑ) $\forall F(w_1 \models Fa \equiv \hat{c}F)$

Our second preliminary fact is obtained by first observing that it is a modally strict theorem that $Ga \equiv \neg\overline{G}a$, by commuting an appropriate instance of (199.2). Hence, by RN, $\Box(Ga \equiv \neg\overline{G}a)$. So by theorem (525):

(ξ) $Ga \Leftrightarrow \neg\overline{G}a$

With (ϑ) and (ξ), we may reason as follows:

$$\begin{array}{llll} \hat{c}G & \equiv & w_1 \models Ga & \text{by } (\vartheta) \text{ and commutativity of} \equiv \\ & \equiv & w_1 \models \neg\overline{G}a & \text{by } (\xi) \text{ and the 1-modal closure of } w_1 \text{ (527)} \\ & \equiv & \neg w_1 \models \overline{G}a & \text{by (529.1)} \\ & \equiv & \neg\hat{c}\overline{G} & \text{by } (\vartheta) \end{array}$$

Since w_1 and a were arbitrarily chosen witnesses, the conclusion $\hat{c}G \equiv \neg\hat{c}\overline{G}$ follows from $\exists w AppearsAt(\hat{c}, w)$. ⋈

(717.2) (Exercise)

(717.3) By (716), we know both:

(ϑ) $\hat{c}G \equiv \hat{c} \succeq \boldsymbol{c}_G$

(ξ) $\hat{c}\overline{G} \equiv \hat{c} \succeq \boldsymbol{c}_{\overline{G}}$

So we may reason as follows:

$$\begin{array}{llll} \hat{c} \succeq \boldsymbol{c}_G & \equiv & \hat{c}G & \text{by } (\vartheta) \text{ and commutativity of} \equiv \\ & \equiv & \neg\hat{c}\overline{G} & \text{by (717.1)} \\ & \equiv & \hat{c} \not\succeq \boldsymbol{c}_{\overline{G}} & \text{by } (\xi) \qquad \bowtie \end{array}$$

(717.4) (Exercise)

(718) By the definition of completeness (687) and GEN, we have to show $\hat{c}F \vee \hat{c}\overline{F}$. Since $\hat{c}$ is a possible-individual concept, we know by the definitions of *PossibleIndividualConcept* (709) and *AppearsAt* (700) that some ordinary object, say a, realizes $\hat{c}$ at some possible world, say w_1. So by definition of *RealizesAt* (697), we know:

(ϑ) $\forall G(w_1 \models Ga \equiv \hat{c}G)$

Now, independently, note that from the easily-proved theorem $\Box(Fa \vee \neg Fa)$ and the instance $\overline{F}a \equiv \neg Fa$ of the modally strict theorem (199.1), it follows by the Rule of Substitution (160.2) that $\Box(Fa \vee \overline{F}a)$. Then by a Fundamental Theorem of Possible World Theory (543.2), it follows that $w_1 \models (Fa \vee \overline{F}a)$. But then by (545.3), it follows that:

(ξ) $w_1 \models Fa \vee w_1 \models \overline{F}a$

So, we may reason by disjunctive syllogism from the disjuncts of (ξ): if $w_1 \models Fa$, then $\hat{c}F$ by (ϑ), and if $w_1 \models \overline{F}a$, then $\hat{c}\overline{F}$, again by (ϑ). Hence, $\hat{c}F \vee \hat{c}\overline{F}$. ⋈

(720) ($\rightarrow$) Assume $\hat{c}$ and $\hat{e}$ are possible-individual concepts that are compossible. Then by definition (719), there is a possible world, say w_1, such that both $AppearsAt(\hat{c}, w_1)$ and $AppearsAt(\hat{e}, w_1)$. But we also know, by (715.1), that $AppearsAt(\hat{c}, \boldsymbol{w}_{\hat{c}})$ and $AppearsAt(\hat{e}, \boldsymbol{w}_{\hat{e}})$. Since possible-individual concepts appear at a unique possible world (713.1), it follows, respectively, that $w_1 = \boldsymbol{w}_{\hat{c}}$ and $w_1 = \boldsymbol{w}_{\hat{e}}$. Hence $\boldsymbol{w}_{\hat{c}} = \boldsymbol{w}_{\hat{e}}$. ($\leftarrow$) Assume $\boldsymbol{w}_{\hat{c}} = \boldsymbol{w}_{\hat{e}}$. Independently, by (715.1), we know both $AppearsAt(\hat{c}, \boldsymbol{w}_{\hat{c}})$ and $AppearsAt(\hat{e}, \boldsymbol{w}_{\hat{e}})$. But from our assumption and the second of these, it follows that $AppearsAt(\hat{e}, \boldsymbol{w}_{\hat{c}})$. But then, there is a world, namely $\boldsymbol{w}_{\hat{c}}$, where both $\hat{c}$ and $\hat{e}$ appear. So $Compossible(\hat{c}, \hat{e})$. ⋈

(721.1) By the definition of *possible-individual concept* (709), $\exists w(Appear(\hat{c}, w)$. Let w_1 be such a world, so that we know $Appear(\hat{c}, w_1)$. By the idempotence of conjunction (85.6), $Appear(\hat{c}, w_1)$ & $Appear(\hat{c}, w_1)$. So by $\exists$I, $\exists w(Appear(\hat{c}, w)$ & $Appear(\hat{c}, w))$. By the definition of compossibility (719), $Compossible(\hat{c}, \hat{c})$. ⋈

(721.2) Suppose $Compossible(\hat{c}, \hat{e})$. Then, by definition (719):

$$\exists w(Appear(\hat{c}, w) \,\&\, Appear(\hat{e}, w))$$

So, by the logic of quantification and the commutativity of conjunction (88.2.a):

$$\exists w(Appear(\hat{e}, w) \,\&\, Appear(\hat{c}, w))$$

Hence, $Compossible(\hat{e}, \hat{c})$, again by the definition of compossibility (719). ⋈

(721.3) Suppose $Compossible(\hat{c}, \hat{d})$ and $Compossible(\hat{d}, \hat{e})$. Then, by a previous theorem (720), $w_{\hat{c}} = w_{\hat{d}}$ and $w_{\hat{d}} = w_{\hat{e}}$. So, by transitivity of identity, $w_{\hat{c}} = w_{\hat{e}}$. Hence, again by (720), $Compossible(\hat{c}, \hat{e})$. ⋈

(723.1) Since $\hat{c}$ is a possible-individual concept, we know, by the definitions of individual concept (709) and appearance (700), that there is an ordinary individual, say b, and a possible world, say w_1, such that $RealizesAt(b, \hat{c}, w_1)$. So, by the idempotence of & (85.6), $RealizesAt(b, \hat{c}, w_1)$ & $RealizesAt(b, \hat{c}, w_1)$. By three applications of $\exists$I, we have $\exists u \exists w \exists w'(RealizesAt(u, \hat{c}, w)$ & $RealizesAt(u, \hat{c}, w'))$. So, by the definition of counterpart (722), $CounterpartOf(\hat{c}, \hat{c})$. ⋈

(723.2) Assume $CounterpartOf(\hat{e}, \hat{c})$. Then, by applying definitions, we know there is an ordinary object, say b, and there are possible worlds, say w_1 and w_2, such that $RealizesAt(b, \hat{c}, w_1)$ & $RealizesAt(b, \hat{e}, w_2)$. By the commutativity of & (88.2.a), $RealizesAt(b, \hat{e}, w_2)$ & $RealizesAt(b, \hat{c}, w_1)$. It follows that:

$$\exists u \exists w \exists w'(RealizesAt(u, \hat{e}, w) \,\&\, RealizesAt(u, \hat{c}, w'))$$

So by the definition of counterparts (722), $CounterpartOf(\hat{c}, \hat{e})$. ⋈

(723.3) Assume $CounterpartOf(\hat{e},\hat{d})$ and $CounterpartOf(\hat{d},\hat{c})$. Then by the first conjunct we know that there is an ordinary object, say a, and there are possible worlds, say w_1 and w_2, such that:

(ϑ) $RealizesAt(a,\hat{d},w_1)\ \&\ RealizesAt(a,\hat{e},w_2)$

And by the second conjunct of our assumption we know that there is an ordinary object, say b, and there are possible worlds, say w_3 and w_4, such that:

(ξ) $RealizesAt(b,\hat{c},w_3)\ \&\ RealizesAt(b,\hat{d},w_4)$

Then, from the first conjunct of (ϑ) and the second conjunct (ξ), it follows by a fact about realization (707) that $w_1 = w_4$ and $a = b$. So after substituting a for b in the first conjunct of (ξ), we may conjoin the result with the second conjunct of (ϑ) to obtain: $RealizesAt(a,\hat{c},w_3)\ \&\ RealizesAt(a,\hat{e},w_2)$. It therefore follows that:

$$\exists u \exists w \exists w'(RealizesAt(u,\hat{c},w)\ \&\ RealizesAt(u,\hat{e},w')),$$

So by definition (722), $CounterpartOf(\hat{e},\hat{c})$. ⋈

(724) ($\rightarrow$) Assume $CounterpartOf(\hat{e},\hat{c})$. Then by the definition of counterpart (722), there is an ordinary object, say a, and there are possible worlds, say, w_1 and w_2, such that:

(ϑ) $RealizesAt(a,\hat{c},w_1)\ \&\ RealizesAt(a,\hat{e},w_2)$

By &I, ∃I, and the definition of the uniqueness quantifier, it remains only to establish:

$$\forall u((RealizesAt(u,\hat{c},w_1)\ \&\ RealizesAt(u,\hat{e},w_2)) \rightarrow u = a)$$

So by GEN, assume:

(ξ) $RealizesAt(u,\hat{c},w_1)\ \&\ RealizesAt(u,\hat{e},w_2)$

But if we conjoin the first conjunct of (ξ) with the first conjunct of (ϑ), we have $RealizesAt(u,\hat{c},w_1)\ \&\ RealizesAt(a,\hat{c},w_1)$. So $u = a$, by a fact about realization (699.2).

($\leftarrow$) Exercise. ⋈

(726.1) – **(726.3)** (Exercises)

(728.1) – **(728.3)** (Exercises)

(729.1) By the second conjunct of (728.3), we know $\forall F(\boldsymbol{c}_u^w F \equiv w \models Fu)$. So, by the commutativity of the biconditional, $\forall F(w \models Fu \equiv \boldsymbol{c}_u^w F)$. Hence, by the definition of realization (697), it follows that $RealizesAt(u,\boldsymbol{c}_u^w,w)$. ⋈

(729.2) By applying GEN to (729.1), we know that $\forall u RealizesAt(u, \boldsymbol{c}_u^w, w)$. But we also know that there are ordinary objects, by (227.1) and the T schema. Hence $\exists u RealizesAt(u, \boldsymbol{c}_u^w, w)$. So, by the definition of appearance (700), it follows that $AppearsAt(\boldsymbol{c}_u^w, w)$. ⋈

(729.3) By applying GEN to (729.2), we know $\forall w AppearsAt(\boldsymbol{c}_u^w, w)$. But we know that there are possible worlds. Hence $\exists w AppearsAt(\boldsymbol{c}_u^w, w)$. So by definition (709), $PossibleIndividualConcept(\boldsymbol{c}_u^w)$. ⋈

(729.4) This follows immediately from (729.2) and (704). ⋈

(729.5) (Exercise)

(730.1) We may reason biconditionally as follows:

$$\begin{array}{lll} PossibleIndividualConcept(c) & \equiv \exists w AppearsAt(c,w) & \text{by definition (709)} \\ & \equiv \exists w \exists u RealizesAt(u,c,w) & \text{by definition (700)} \\ & \equiv \exists u \exists w RealizesAt(u,c,w) & \text{by theorem (103.11)} \\ & \equiv \exists u \exists w \forall F(w \models Fu \equiv cF) & \text{by definition (697)} \\ & \equiv \exists u \exists w \forall F(cF \equiv w \models Fu) & \text{(88.2.e) and (159.3)} \\ & \equiv \exists u \exists w ConceptOfAt(c,u,w) & \text{by definition (725)} \bowtie \end{array}$$

(730.2) (→) Assume $PossibleIndividualConcept(c)$. Then by definition (709), we know $\exists w AppearsAt(c,w)$. Suppose w_1 is an arbitrary such world, so that we have $AppearsAt(c,w_1)$. By the definition of appearance (700), some ordinary individual, say b, is such that $RealizesAt(b,c,w_1)$. So $\forall F(w_1 \models Fb \equiv cF)$, by the definition of realization (697). Now the second conjunct of (728.3) tells us that $\forall F(\boldsymbol{c}_b^{w_1} F \equiv w_1 \models Fb)$. So by the logic of quantified biconditionals, we know: $\forall F(cF \equiv \boldsymbol{c}_b^{w_1} F)$. Since both c and $\boldsymbol{c}_b^{w_1}$ are concepts and, hence, abstract, it follows that $c = \boldsymbol{c}_b^{w_1}$. So by ∃I, $\exists u \exists w(c = \boldsymbol{c}_u^w)$.

(←) Assume $\exists u \exists w(c = \boldsymbol{c}_u^w)$, to show $PossibleIndividualConcept(c)$. So $c = \boldsymbol{c}_b^{w_1}$, for some arbitrary ordinary object b and possible world w_1. Now we independently know that $IndividualConcept(\boldsymbol{c}_b^{w_1})$, by (729.3). So $IndividualConcept(c)$. ⋈

(731) In Remark (684), we noted that c_u can be canonically identified:

$$\boldsymbol{c}_u = \iota x(A!x \,\&\, \forall F(xF \equiv Fu))$$

Moreover, as an instance of Actualized Abstraction (258.2), we know:

$$\iota x(A!x \,\&\, \forall F(xF \equiv Fu))G \equiv \mathcal{A}Gu$$

From these two results, it follows that:

$$\boldsymbol{c}_u G \equiv \mathcal{A}Gu$$

But as an instance of (538), we know:

$$\mathcal{A}Gu \equiv \boldsymbol{w}_\alpha \models Gu$$

Our last two results yield $\boldsymbol{c}_u G \equiv \boldsymbol{w}_\alpha \models Gu$, which by the commutativity of the biconditional leaves us with:

(ϑ) $\boldsymbol{w}_\alpha \models Gu \equiv \boldsymbol{c}_u G$

Now, independently, since $\boldsymbol{w}_\alpha$ is a possible world (537.1), it follows from the second conjunct of (728.3) that:

(ζ) $\boldsymbol{c}_u^{\boldsymbol{w}_\alpha} G \equiv \boldsymbol{w}_\alpha \models Gu$

(ζ) and (ϑ) imply $\boldsymbol{c}_u^{\boldsymbol{w}_\alpha} G \equiv \boldsymbol{c}_u G$. So by GEN, $\forall G(\boldsymbol{c}_u^{\boldsymbol{w}_\alpha} G \equiv \boldsymbol{c}_u G)$. Since both $\boldsymbol{c}_u^{\boldsymbol{w}_\alpha}$ and $\boldsymbol{c}_u$ are abstract, it follows that $\boldsymbol{c}_u^{\boldsymbol{w}_\alpha} = \boldsymbol{c}_u$. ⋈

(732.1) By (729.3) $\boldsymbol{c}_u^w$ is a possible-individual concept. So by (716), it follows that $\boldsymbol{c}_u^w G \equiv \boldsymbol{c}_u^w \succeq \boldsymbol{c}_G$. ⋈

(732.2) From the second conjunct of (728.3), it follows that $\boldsymbol{c}_u^w G \equiv w \models Gu$, and so by commutativity of the biconditional, $(w \models Gu) \equiv \boldsymbol{c}_u^w G$. But we just established (732.1) that $\boldsymbol{c}_u^w G \equiv \boldsymbol{c}_u^w \succeq \boldsymbol{c}_G$. Hence, $(w \models Gu) \equiv \boldsymbol{c}_u^w \succeq \boldsymbol{c}_G$. ⋈

(732.3) Assume $\boldsymbol{c}_u^w = \boldsymbol{c}_v^w$. Independently, we know both $RealizesAt(u, \boldsymbol{c}_u^w, w)$ and $RealizesAt(v, \boldsymbol{c}_v^w, w)$, by (729.1). So we may substitute $\boldsymbol{c}_u^w$ for $\boldsymbol{c}_v^w$ in the latter, to obtain $RealizesAt(v, \boldsymbol{c}_u^w, w)$. But from $RealizesAt(u, \boldsymbol{c}_u^w, w)$ and $RealizesAt(v, \boldsymbol{c}_u^w, w)$, it follows from a fact about realization (699.2), that $u = v$. ⋈

(732.4) Assume $\boldsymbol{c}_u^w = \boldsymbol{c}_u^{w'}$. Independently, we know both $RealizesAt(u, \boldsymbol{c}_u^w, w)$ and $RealizesAt(u, \boldsymbol{c}_u^{w'}, w')$, by (729.1). So we may substitute $\boldsymbol{c}_u^w$ for $\boldsymbol{c}_u^{w'}$ in the latter, to infer $RealizesAt(u, \boldsymbol{c}_u^w, w')$. But from $RealizesAt(u, \boldsymbol{c}_u^w, w)$ and $RealizesAt(u, \boldsymbol{c}_u^w, w')$, it follows from a fact about realization (699.3), that $w = w'$. ⋈

(732.5) (Exercise)

(733.1) From a fact about world-relative concepts of individuals (729.1), we know both $RealizesAt(u, \boldsymbol{c}_u^w, w)$ and $RealizesAt(v, \boldsymbol{c}_v^w, w)$. Since ordinary objects exist, by (227.1) and the T schema, we may infer, respectively:

$$\exists u RealizesAt(u, \boldsymbol{c}_u^w, w)$$
$$\exists v RealizesAt(v, \boldsymbol{c}_v^w, w)$$

So, by definition (700), it follows, respectively, that:

$$AppearsAt(\boldsymbol{c}_u^w, w)$$
$$AppearsAt(\boldsymbol{c}_v^w, w)$$

By conjoining these results and quantifying:

$$\exists w'(AppearsAt(\boldsymbol{c}_u^w, w') \,\&\, AppearsAt(\boldsymbol{c}_v^w, w'))$$

So by the definition of compossibility (719), $Compossible(\boldsymbol{c}_u^w, \boldsymbol{c}_v^w)$. ⋈

(733.2) (Exercise)

(733.3) By (729.1), we know both $RealizesAt(u, \boldsymbol{c}_u^w, w)$ and $RealizesAt(u, \boldsymbol{c}_u^{w'}, w')$, which we may conjoin. Since the restricted variables u, w, and w' all range over non-empty domains, we obtain by ∃I:

$$\exists u \exists w \exists w'(RealizesAt(u, \boldsymbol{c}_u^w, w) \,\&\, RealizesAt(u, \boldsymbol{c}_u^{w'}, w'))$$

So, by the definition of counterparts (722), it follows that $CounterpartOf(\boldsymbol{c}_u^{w'}, \boldsymbol{c}_u^w)$. ⋈

(733.4) (Exercise)

(736.1)★ Assume:

(ϑ) $Fu \,\&\, \Diamond\neg Fu$

The first conjunct and theorem (692.1)★, which asserts $Gu \equiv \boldsymbol{c}_u \succeq \boldsymbol{c}_G$, together imply $\boldsymbol{c}_u \succeq \boldsymbol{c}_F$. So it remains show:

(ξ) $\exists \hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u) \,\&\, \hat{c} \not\succeq \boldsymbol{c}_F \,\&\, \exists w(w \neq \boldsymbol{w}_\alpha \,\&\, AppearsAt(\hat{c}, w)))$

Note that from the second conjunct of (ϑ) and a Fundamental Theorem of Possible World Theory (543.1), it follows that $\exists w(w \models \neg Fu)$. So, let w_1 be an arbitrary such possible world, so that we know $w_1 \models \neg Fu$. Now consider the concept of u at w_1, i.e., $\boldsymbol{c}_u^{w_1}$. We know by (729.3) that $PossibleIndividualConcept(\boldsymbol{c}_u^{w_1})$. So to show ($\xi$), it suffices by &I and ∃I to show:

(a) $CounterpartOf(\boldsymbol{c}_u^{w_1}, \boldsymbol{c}_u)$

(b) $\boldsymbol{c}_u^{w_1} \not\succeq \boldsymbol{c}_F$

(c) $\exists w(w \neq \boldsymbol{w}_\alpha \,\&\, AppearsAt(\boldsymbol{c}_u^{w_1}, w))$

(a) If we instantiate theorem (733.3) to worlds w_1 and $\boldsymbol{w}_\alpha$, it follows that:

$$CounterpartOf(\boldsymbol{c}_u^{w_1}, \boldsymbol{c}_u^{\boldsymbol{w}_\alpha})$$

But by (731), we know $\boldsymbol{c}_u^{\boldsymbol{w}_\alpha} = \boldsymbol{c}_u$. So it follows that $CounterpartOf(\boldsymbol{c}_u^{w_1}, \boldsymbol{c}_u)$.
(b) Since we know $w_1 \models \neg Fu$, it follows by (529.1) that $\neg w_1 \models Fu$. So by the second conjunct of (728.3), it follows that $\neg \boldsymbol{c}_u^{w_1} F$. But since $\boldsymbol{c}_u^{w_1}$ is known to be a possible-individual concept, it then follows from (732.1) that $\boldsymbol{c}_u^{w_1} \not\succeq \boldsymbol{c}_F$.
(c) By &I and ∃I, it suffices to show both $w_1 \neq \boldsymbol{w}_\alpha$ and $AppearsAt(\boldsymbol{c}_u^{w_1}, w_1)$. Note that by a theorem of possible world theory (536.2)★, the first conjunct of (ϑ) is equivalent to $\boldsymbol{w}_\alpha \models Fu$. But given this last fact and the previously established fact that $\neg w_1 \models Fu$, there is a proposition, namely Fu, that is true at $\boldsymbol{w}_\alpha$ but not true at w_1. Since worlds are situations, it follows by (474) that $w_1 \neq \boldsymbol{w}_\alpha$. So

it remains to show $AppearsAt(\boldsymbol{c}_u^{w_1}, w_1)$. But this is immediate from a fact about appearance and world-relative concepts of individuals (729.2). ⋈

(736.2)★ (Exercise)

(737.1)★ (→) This direction follows *a fortiori* from the left-to-right direction of (736.1)★. (←) Assume the antecedent:

$$\boldsymbol{c}_u \geq \boldsymbol{c}_F \ \& \ \exists \hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u) \ \& \ \hat{c} \not\geq \boldsymbol{c}_F)$$

From the first conjunct and (692.1)★, it follows that Fu. So it remains to show $\Diamond\neg Fu$. By a fundamental theorem of possible worlds (543.1), it suffices to show $\exists w(w \models \neg Fu)$. So our proof strategy is to find a witness to this latter claim.

Suppose $\hat{c}_1$ is a possible-individual concept that witnesses the second conjunct of our assumption, so that we know:

(ϑ) $CounterpartOf(\hat{c}_1, \boldsymbol{c}_u) \ \& \ \hat{c}_1 \not\geq \boldsymbol{c}_F$

Now the first conjunct of (ϑ) implies by definition (722) that:

$$\exists v \exists w \exists w'(RealizesAt(v, \boldsymbol{c}_u, w) \ \& \ RealizesAt(v, \hat{c}_1, w'))$$

Assume ordinary object b and possible worlds w_1 and w_2 are witnesses to this existential claim. Then we know:

(ξ) $RealizesAt(b, \boldsymbol{c}_u, w_1) \ \& \ RealizesAt(b, \hat{c}_1, w_2)$

Now independently by (708.1), we know $RealizesAt(u, \boldsymbol{c}_u, \boldsymbol{w}_\alpha)$. This and the first conjunct of (ξ) imply, by (707), that $w_1 = \boldsymbol{w}_\alpha$ and $b = u$. From the latter and the second conjunct of (ξ) it follows that $RealizesAt(u, \hat{c}_1, w_2)$. By definition (697), this implies $\forall G(w_2 \models Gu \equiv \hat{c}_1 G)$. So, in particular, $w_2 \models Fu \equiv \hat{c}_1 F$. Now assume, for reductio, that $w_2 \models Fu$. Then $\hat{c}_1 F$. But this implies $\hat{c}_1 \geq \boldsymbol{c}_F$, by (716), which contradicts the second conjunct of (ϑ). Hence $\neg(w_2 \models Fu)$, and so by the coherence of truth at a world (529.1), it follows that $w_2 \models \neg Fu$. Thus, w_2 is a witness to $\exists w(w \models \neg Fu)$, which is what we had to find. ⋈

(737.2)★ (Exercise)

(738.1) Assume $w \models (Fu \ \& \ \Diamond\neg Fu)$. Then since the laws of conjunction hold with respect to truth at a possible world (545.1), it follows that:

(ϑ) $w \models Fu \ \& \ w \models \Diamond\neg Fu$

Now the first conjunct of (ϑ) and the second conjunct of theorem (728.3), which asserts $\forall G(\boldsymbol{c}_u^w G \equiv w \models Gu)$, together imply $\boldsymbol{c}_u^w F$. This implies, by (732.1), that $\boldsymbol{c}_u^w \geq \boldsymbol{c}_F$. We've therefore established the first conjunct of our desired conclusion. So it remains show:

(ξ) $\exists \hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^w) \ \& \ \hat{c} \not\geq \boldsymbol{c}_F \ \& \ \exists w'(w' \neq w \ \& \ AppearsAt(\hat{c}, w')))$

Note that from the second conjunct of (ϑ), it follows that $\exists w'(w' \models \Diamond\neg Fu)$. So by a Fundamental Theorem of Possible World Theory (543.1), it follows that $\Diamond\Diamond\neg Fu$. This implies $\Diamond\neg Fu$, by the 4$\Diamond$ schema (165.7). So by the same Fundamental Theorem, $\exists w'(w' \models \neg Fu)$. Now let w_1 be an arbitrary such possible world, so that we know $w_1 \models \neg Fu$, and consider the concept of u at w_1, i.e., $\boldsymbol{c}_u^{w_1}$. We know by (729.3) that $PossibleIndividualConcept(\boldsymbol{c}_u^{w_1})$. So to show ($\xi$), it suffices by &I and $\exists$I to show:

(a) $CounterpartOf(\boldsymbol{c}_u^{w_1}, \boldsymbol{c}_u^{w})$

(b) $\boldsymbol{c}_u^{w_1} \not\succeq \boldsymbol{c}_F$

(c) $\exists w'(w' \neq w \;\&\; AppearsAt(\boldsymbol{c}_u^{w_1}, w'))$

(a) This is an instance of theorem (733.3).

(b) Since we know $w_1 \models \neg Fu$, it follows by (529.1) that $\neg w_1 \models Fu$. So by the second conjunct of (728.3), it follows that $\neg\boldsymbol{c}_u^{w_1}F$. But since $\boldsymbol{c}_u^{w_1}$ is known to be a possible-individual concept, it then follows from (732.1) that $\boldsymbol{c}_u^{w_1} \not\succeq \boldsymbol{c}_F$.

(c) By &I and $\exists$I, it suffices to show both $w_1 \neq w$ and $AppearsAt(\boldsymbol{c}_u^{w_1}, w_1)$. By the first conjunct of (ϑ), we know $w \models Fu$. But we've previously established $\neg w_1 \models Fu$. So there is a proposition, namely Fu, that is true at w but not true at w_1. Since worlds are situations, it follows by (474) that $w_1 \neq w$. So it remains to show $AppearsAt(\boldsymbol{c}_u^{w_1}, w_1)$. But this is an instance of a fact about appearance and world-relative concepts of individuals (729.2). ⋈

(738.2) (Exercise)

(738.3) ($\rightarrow$) This direction follows *a fortiori* from the left-to-right direction of (738.1). ($\leftarrow$) Assume the antecedent:

$$\boldsymbol{c}_u^{w} \succeq \boldsymbol{c}_F \;\&\; \exists\hat{c}(CounterpartOf(\hat{c}, \boldsymbol{c}_u^{w}) \;\&\; \hat{c} \not\succeq \boldsymbol{c}_F)$$

To show $w \models (Fu \;\&\; \Diamond\neg Fu)$, it suffices to show both $w \models Fu$ and $w \models \Diamond\neg Fu$, in virtue of the right-to-left condition of (545.1). But from the first conjunct of our assumption and (732.2), it follows that $w \models Fu$. So it remains to show $w \models \Diamond\neg Fu$. Suppose $\hat{c}_1$ is a witness to the second conjunct of our assumption, so that we know:

(ϑ) $CounterpartOf(\hat{c}_1, \boldsymbol{c}_u^{w}) \;\&\; \hat{c}_1 \not\succeq \boldsymbol{c}_F$

Now the first conjunct of (ϑ) implies by definition (722) that:

$$\exists v \exists w' \exists w''(RealizesAt(v, \boldsymbol{c}_u^{w}, w') \;\&\; RealizesAt(v, \hat{c}_1, w''))$$

Assume ordinary object b and possible worlds w_1 and w_2 are witnesses to this existential claim. Then we know:

(ξ) $RealizesAt(b, \boldsymbol{c}_u^w, w_1)\,\&\,RealizesAt(b, \hat{c}_1, w_2)$

Now independently by (729.1), we know $RealizesAt(u, \boldsymbol{c}_u^w, w)$. This and the first conjunct of (ξ) imply, by (707), that $w_1 = w$ and $b = u$. From the latter and the second conjunct of (ξ) it follows that $RealizesAt(u, \hat{c}_1, w_2)$. By definition (697), this implies $\forall G(w_2 \models Gu \equiv \hat{c}_1 G)$. So, in particular, $w_2 \models Fu \equiv \hat{c}_1 F$. Now assume, for reductio, that $w_2 \models Fu$. Then $\hat{c}_1 F$. But this implies $\hat{c}_1 \succeq \boldsymbol{c}_F$, by (716), which contradicts the second conjunct of (ϑ). Hence $\neg(w_2 \models Fu)$, and so by the coherence of truth at a world (529.1), it follows that $w_2 \models \neg Fu$. So $\exists w(w \models \neg Fu)$ and, hence, by a fundamental theorem of possible worlds (543.1), $\Diamond\neg Fu$. Then by the 5 schema (45.3), $\Box\Diamond\neg Fu$, and so by another fundamental theorem of possible worlds (543.2), $\forall w'(w' \models \Diamond\neg Fu)$. But then $w \models \Diamond\neg Fu$. $\bowtie$

(738.4) (Exercise)

(744) In the usual manner, we need not concern ourselves with the existence clauses in the definientia of the notions involved. We are citing only definitions instanced to variables, and these instances imply biconditionals between a definiendum and its definiens without the existence claims.

($\rightarrow$) Assume $R|\!: F \xleftrightarrow{\text{1-1}} G$. Then by (741) we know both:

(A) $\forall x(Fx \rightarrow \exists!y(Gy\,\&\,Rxy))$

(B) $\forall y(Gy \rightarrow \exists!x(Fx\,\&\,Rxy))$

By (743.4), we have to show both:

(a) $R|\!: F \xrightarrow{\text{1-1}} G$

(b) $R|\!: F \xrightarrow[\text{onto}]{} G$

(a) By (743.2), we have to show both:

$R|\!: F \longrightarrow G$

$\forall x\forall y\forall z((Fx\,\&\,Fy\,\&\,Gz) \rightarrow (Rxz\,\&\,Ryz \rightarrow x = y))$

By (743.1), the first just means $\forall x(Fx \rightarrow \exists!y(Gy\,\&\,Rxy))$, which is just (A). So it remains to show the second. Assume $Fx\,\&\,Fy\,\&\,Gz$. And further assume $Rxz\,\&\,Ryz$. Then from Gz and (B), it follows that there is a unique x that bears R to z. Then from Fx, Fy, Rxz, and Ryz it follows that $x = y$, on pain of contradiction.

(b) By (743.3), we have to show both:

$R|\!: F \longrightarrow G$

$\forall y(Gy \rightarrow \exists x(Fx\,\&\,Rxy))$

But we've seen that the first is already known. And the second follows *a fortiori* from (B).

(←) Assume $R|\!:\! F \xrightarrow[\text{onto}]{\text{1-1}} G$. After applying definitions (743.4), (743.2), and (743.3), and then simplifying, we know:

(C) $R|\!:\! F \longrightarrow G$

(D) $\forall x \forall y \forall z((Fx \,\&\, Fy \,\&\, Gz) \to (Rxz \,\&\, Ryz \to x = y))$

(E) $\forall y(Gy \to \exists x(Fx \,\&\, Rxy))$

We have to show:

(c) $\forall x(Fx \to \exists! y(Gy \,\&\, Rxy))$

(d) $\forall y(Gy \to \exists! x(Fx \,\&\, Rxy))$

(c) This follows from (C) and definition (743.1).

(d) By GEN, it suffices to show $Gy \to \exists! x(Fx \,\&\, Rxy)$. So assume Gy. By (E), it follows that $\exists x(Fx \,\&\, Rxy)$. So suppose a is such an object, so that we know Fa and Ray. Then by &I, ∃I and the definition of the uniqueness quantifier, it remains only to show $\forall z(Fz \,\&\, Rzy \to z = a)$. So assume $Fz \,\&\, Rzy$. But since we now know Fz, Fa, Gy, Rzy, and Ray, we may infer $z = a$ by (D). ⋈

(747.1) Pick a variable, say z, that is substitutable for x in φ. Then if we eliminate the restricted variable, we have to show:

$$\exists! x(D!x \,\&\, \varphi) \equiv \exists x(D!x \,\&\, \varphi \,\&\, \forall z((D!z \,\&\, \varphi^z_x) \to z = x))$$

But, clearly, this is an instance of the definition of the uniqueness quantifier (127). ⋈

(748.1) Given Rule ≡S and the facts that $R\!\downarrow$, $F\!\downarrow$, and $G\!\downarrow$, it suffices by definitions (747.3) and (747.2) to show that:

$$\exists R[\forall u(Fu \to \exists! v(Fv \,\&\, Ruv)) \,\&\, \forall v(Fv \to \exists! u(Fu \,\&\, Ruv))]$$

If we propose $=_D$ as our witness, then we have to show:

(a) $\forall u(Fu \to \exists! v(Fv \,\&\, u =_D v))$

(b) $\forall v(Fv \to \exists! u(Fu \,\&\, u =_D v))]$

(a) By GEN, it suffices to show $Fu \to \exists! v(Fv \,\&\, u =_D v)$. So assume Fu. We have to show, by the definition of the unique existence quantifier for discernible objects (747.1):

$$\exists v(Fv \,\&\, u =_D v \,\&\, \forall v'((Fv' \,\&\, u =_D v') \to v' = v))$$

But if we propose that u is a witness to this claim, then we have to show:

$$Fu \mathbin{\&} u =_D u \mathbin{\&} \forall v'((Fv' \mathbin{\&} u =_D v') \rightarrow v' = u)$$

We know Fu by assumption, and $u =_D u$ is a theorem. So it remains only to show:

$$\forall v'((Fv' \mathbin{\&} u =_D v') \rightarrow v' = u)$$

But this is trivial, by the fact that $u =_D v' \rightarrow u = v'$ (273.19) and the symmetry of =.

(b) Note that the following is an alphabetic variant of what we have to show:

$$\forall u(Fu \rightarrow \exists! v(Fv \mathbin{\&} v =_D u))$$

We leave it as an exercise to show that this follows from (a) by a modally strict derivation that appeals to the symmetry of $=_D$ (273.31). ⋈

(748.2) To show that equinumerosity$_D$ is symmetric, assume that $F \approx_D G$ and suppose R is a witness to this fact. Then by Rule ≡S and the facts that $R\!\downarrow$, $F\!\downarrow$, and $G\!\downarrow$, it follows by definitions (747.3) and (747.2) that:

$$(\vartheta)\quad \forall u(Fu \rightarrow \exists! v(Gv \mathbin{\&} Ruv)) \mathbin{\&} \forall v(Gv \rightarrow \exists! u(Fu \mathbin{\&} Ruv))$$

Now we want to show, for some R', that $R'|\!: G \xrightarrow{1\text{-}1}_{\text{onto}} F$. Consider the converse of R, namely $[\lambda xy\, Ryx]$. Clearly, $[\lambda xy\, Ryx]\!\downarrow$, by (39.2). So let us call this property R^{-1}. Clearly, $R^{-1}\!\downarrow$. By Rule ≡S, the facts that $R\!\downarrow$, $F\!\downarrow$, and $G\!\downarrow$, and definitions (747.3) and (747.2), we need to show:

(a) $\forall u(Gu \rightarrow \exists! v(Fv \mathbin{\&} R^{-1}uv))$

(b) $\forall v(Fv \rightarrow \exists! u(Gu \mathbin{\&} R^{-1}vu))$

(a) Assume Gu, by GEN. To avoid clash of variables, note that the second conjunct of (ϑ) is equivalent to the following alphabetic variant:

$$\forall v(Gv \rightarrow \exists! u'(Fu' \mathbin{\&} Ru'v))$$

So, if instantiate u into this claim, we can, given our assumption Gu, detach the consequent from the result, to conclude $\exists! u'(Fu' \mathbin{\&} Ru'u)$, i.e., again by alphabetic variance, $\exists! v(Fv \mathbin{\&} Rvu)$. Now, independently, since we know that $R^{-1}\!\downarrow$ by definition, β-Conversion and the definition of R^{-1} imply the following modally strict equivalence:

$$R^{-1}uv \equiv Rvu$$

So, by the Rule of Substitution (160.2), $\exists!v(Fv \,\&\, R^{-1}uv)$.

(b) By analogous reasoning. ⋈

(748.3) To show that equinumerosity$_D$ is transitive, assume both that $F \approx_D G$ and $G \approx_D H$. Suppose R_1 and R_2 are relations that bear witness to these facts, respectively. Then by Rule ≡S and the facts that $R\downarrow$, $F\downarrow$, and $G\downarrow$, it follows from definitions (747.3) and (747.2) that:

(ϑ) $\forall u(Fu \to \exists!v(Gv \,\&\, R_1uv)) \,\&\, \forall v(Gv \to \exists!u(Fu \,\&\, R_1uv))$

(ξ) $\forall u(Gu \to \exists!v(Hv \,\&\, R_2uv)) \,\&\, \forall v(Hv \to \exists!u(Gu \,\&\, R_2uv))$

Now let r and s be additional restricted variables ranging over discernible objects and consider the relation:

$$[\lambda rs\, \exists v(Gv \,\&\, R_1rv \,\&\, R_2vs)]$$

Clearly, $[\lambda rs\, \exists v(Gv \,\&\, R_1rv \,\&\, R_2vs)]\downarrow$, so let us call it R.[477] By now familiar reasoning and definitions (747.3) and (747.2), to show that $R|: F \xleftrightarrow{\text{1-1}} H$, we must show:

(a) $\forall u(Fu \to \exists!v(Hv \,\&\, Ruv))$

(b) $\forall v(Hv \to \exists!u(Fu \,\&\, Ruv))$

To show (a), assume Fu, by GEN. By the first conjunct of (ϑ) and (747.1), there is a discernible object, say b, such that both:

(c) $Gb \,\&\, R_1ub$

(d) $\forall v((Gv \,\&\, R_1uv) \to v = b)$

Then, after instantianting b into the first conjunct of (ξ), we know, given the first conjunct of (c), that there is a discernible object, say c, such that both:

(e) $Hc \,\&\, R_2bc$

(f) $\forall v((Hv \,\&\, R_2bv) \to v = c)$

Now to prove $\exists!v(Hv \,\&\, Ruv)$, we choose c as our witness. It then suffices, by ∃I, the definition of the uniqueness quantifier for discernible objects (747.1), and the fact that c is discernible, to show that:

[477] By our conventions for bound occurrences of restricted variables, the λ-expression R abbreviates in turn abbreviates:

$$[\lambda xy\, D!x \,\&\, D!y \,\&\, \exists z(D!z \,\&\, Gz \,\&\, R_1xz \,\&\, R_2zy)]$$

By (337.4), the above eliminates bound occurrences of r and s with bound occurrences of x and y, and by (337.2), the above eliminates bound occurrences of v with bound occurrences of z.

(g) $Hc \,\&Ruc$

(h) $\forall v((Hv \,\&\, Ruv) \rightarrow v = c)$

To show (g), note that Hc is the first conjunct of (e). Now by the fact that $R\downarrow$, the definition of R and β-Conversion, we know Ruc iff $\exists v(Gv \,\&\, R_1uv \,\&\, R_2vc)$. But (c) and (e) above establish that b is witness to the existence claim. To show (h), assume $Hv \,\&\, Ruv$, to show $v = c$. Since Ruv, we know, by the existence and definition of R and β-Conversion, that there is a discernible object, say e, such that $Ge \,\&\, R_1ue \,\&\, R_2ev$. But the first two conjuncts imply, by (d), that $e = b$. This last fact and R_2ev imply R_2bv. So we know $Hv \,\&\, R_2bv$. Then by (f), it follows that $v = c$.

To show (b), assume Hv. Then by the 2nd conjunct of (ξ), we know, for some discernible object, say c:

(i) $Gc \,\&\, R_2cv \,\&\, \forall u((Gu \,\&\, R_2uv) \rightarrow u = c)$

Now if we instantiate c into the second conjunct of (ϑ), then from the first conjunct of (i), we know that for some discernible object, say d:

(j) $Fd \,\&\, R_1dc \,\&\, \forall u((Fu \,\&\, R_1uc) \rightarrow u = d)$

To complete the proof of (b), it suffices, by $\exists$I and the definition of the uniqueness quantifier, to show:

$$Fd \,\&\, Rdv \,\&\, \forall u((Fu \,\&\, Ruv) \rightarrow u = d)$$

Fd is the first conjunct of (j). Moreover, by β-Conversion and the definition and existence of R, we know Rdv, since there is a discernible object, namely c, such that Gc, R_1dc, and R_2cv, by (i) and (j). Since it now remains only to show $\forall u((Fu \,\&\, Ruv) \rightarrow u = d)$, assume $Fu \,\&\, Ruv$. From Ruv and the definition of R, it follows that some discernible object, say e, is such that $Ge \,\&\, R_1ue \,\&\, R_2ev$. From Ge, R_2ev, and the third conjunct of (i), it follows that $e = c$. But then the third conjunct of (j) implies $\forall u((Fu \,\&\, R_1ue) \rightarrow u = d)$. So by $\forall$E, $(Fu \,\&\, R_1ue) \rightarrow u = d$. But we already know both Fu and R_1ue. Hence $u = d$. ⋈

(748.4) ($\rightarrow$) Assume $F \approx_D G$. Then by GEN, it suffices to show $H \approx_D F \equiv H \approx_D G$:

($\rightarrow$) Assume $H \approx_D F$, then by transitivity (748.3), $H \approx_D G$.

($\leftarrow$) Assume $H \approx_D G$. Then, by symmetry (748.2), our initial assumption implies $G \approx_D F$. So by transitivity, $H \approx_D F$.

($\leftarrow$) Assume $\forall H(H \approx_D F \equiv H \approx_D G)$. By instantiating to F we obtain: $F \approx_D F \equiv F \approx_D G$. So by reflexivity (748.1), $F \approx_D G$. ⋈

(750) (Exercise)

(751.1) Assume $\neg\exists uFu$ and $\neg\exists vHv$. Then pick any relation R you please. By failures of the antecedent, we know both:

$$Fu \to \exists!v(Hv \,\&\, Ruv)$$

$$Hv \to \exists!u(Fu \,\&\, Ruv)$$

So by GEN:

$$\forall u(Fu \to \exists!v(Hv \,\&\, Ruv))$$

$$\forall v(Hv \to \exists!u(Fu \,\&\, Ruv))$$

Hence, by &I and definition (747.2) (as simplified by Rule ≡S and the facts that $R\downarrow$, $F\downarrow$, and $H\downarrow$), it follows $R|: F \xleftrightarrow{1\text{-}1}_D H$. So by ∃I and definition (747.3), $F \approx_D H$. ⋈

(751.2) Assume $\exists uFu$ and $\neg\exists vHv$. Given the former, suppose b is such a discernible object, so that we know Fb. Now assume, for reductio, that $F \approx_D H$. Then by definitions (747.3) and (747.2), as simplified by Rule ≡S and the facts that $R\downarrow$, $F\downarrow$, and $H\downarrow$, it follows that:

$$(\vartheta)\;\; \exists R[\forall u(Fu \to \exists!v(Hv \,\&\, Ruv)) \,\&\, \forall v(Hv \to \exists!u(Fu \,\&\, Ruv))]$$

Suppose R_1 is such an R. Then from the first conjunct of (ϑ) and Fb, it follows that $\exists!v(Hv \,\&\, Ruv)$. *A fortiori*, $\exists vHv$. Contradiction. ⋈

(752.1) (Exercise)

(753) Since the following proof is conducted entirely using variables and constants, all of which are significant, we may omit any existence clauses that would otherwise be required by the definitions of the notions involved. Let us use r, s, and t as additional restricted variables for discernible objects, so that r, s, t, u, v all range over discernible objects. Assume that $F \approx_D G$, Fu, and Gv. The first assumption implies, by definition (747.3), that some relation, say R, correlates$_D$ F and G one-to-one, i.e., by definition (747.2):

Fact 1: $\forall r(Fr \to \exists!s(Gs \,\&\, Rrs)) \,\&\, \forall s(Gs \to \exists!r(Fr \,\&\, Rrs))$

Moreover, by (750), it follows that R maps$_D$ F onto G one-to-one. By definition (749.2), the one-to-one character of R entails:

Fact 2: $\forall r\forall s\forall t((Fr \,\&\, Fs \,\&\, Gt) \to (Rrt \,\&\, Rst \to r = s))$

Now we want to show that $F^{-u} \approx_D G^{-v}$. By definitions (747.3) and (747.2), we have to show:

Claim: There is a relation R' such that:

(A) $\forall r(F^{-u}r \to \exists!s(G^{-v}s \,\&\, R'rs))$

(B) $\forall s(G^{-v}s \to \exists!r(F^{-u}r \,\&\, R'rs))$

We proceed by showing that if Ruv (**Case 1**), then R itself is the witness to the **Claim** and if $\neg Ruv$ (**Case 2**), then there exists a relation that can be defined in terms of R that serves as a witness to the **Claim**.

Case 1: Ruv. Then we show R itself is the witness to the **Claim**, i.e., that both (A) and (B) hold with respect to R. In what follows, we use a, b, c, d as constants for discernible objects.

(A) By GEN, suppose $F^{-u}r$, to show $\exists!s(G^{-v}s \,\&\, Rrs)$). Then Fr and $r \neq u$, by the definition of F^{-u} (752.2). But since Fr, the first conjunct of **Fact 1** implies $\exists!s(Gs \,\&\, Rrs)$. Let b be such an object, so that we know:

(ζ) $Gb \,\&\, Rrb \,\&\, \forall t(Gt \,\&\, Rrt \rightarrow t = b)$

We can now show, using b as a witness, that $\exists!s(G^{-v}s \,\&\, Rrs)$. Since we now know Fr, Fu, Gb, Rrb, Ruv, and $r \neq u$, it follows by **Fact 2** that $b \neq v$, on pain of contradiction. Since we have that Gb and $b \neq v$, it follows that $G^{-v}b$. Since we've established $G^{-v}b$ and Rrb, it remains to show uniqueness. By GEN, it suffices to show $(G^{-v}t \,\&\, Rrt) \rightarrow t = b$. So suppose $G^{-v}t$ and Rrt. Then by definition of G^{-v}, it follows that Gt. But then $t = b$, by the last conjunct of (ζ).

(B) By GEN, it suffices to show $G^{-v}s \rightarrow \exists!r(F^{-u}r \,\&\, Rrs)$. So assume $G^{-v}s$. Then, by definition of G^{-v} (752.2), we know Gs and $s \neq v$. From Gs and the second conjunct of **Fact 1**, we know $\exists!r(Fr \,\&\, Rrs)$. So suppose a is such an object, so that we know:

(ω) $Fa \,\&\, Ras \,\&\, \forall t(Ft \,\&\, Rts \rightarrow t = a)$

We can now show, using a as a witness, that $\exists!r(F^{-u}r \,\&\, Rrs)$. Since we now know Fa, Gs, Gv, Ras, Ruv, and $s \neq v$, it follows from the fact that R maps$_D$ F to G (i.e., the first conjunct of **Fact 1**), that $a \neq u$, on pain of contradiction (if a were identical to u, we would have Ras, Rav, and $s \neq v$, which would contradict the fact that R relates a to a unique discernible object exemplifying G). Given we know Fa and $a \neq u$, we have $F^{-u}a$. So we have established $F^{-u}a$ and Ras. It then remains to prove uniqueness. By GEN, it suffices to show $(F^{-u}t \,\&\, Rts) \rightarrow t = a$. So suppose $F^{-u}t$ and Rts. Then Ft, by definition of F^{-u}. So by the last conjunct of (ω), $t = a$.

Case 2: $\neg Ruv$. Since we've assumed Fu and Gv, we therefore know by **Fact 1** both:

> $\exists!s(Gs \,\&\, Rus)$, i.e., there is a unique discernible object that exemplifies G to which u bears R, and
>
> $\exists!r(Fr \,\&\, Rrv)$, i.e., there is a unique discernible object that exemplifies F and that bears R to v.

Let b be a witness to the first and a be a witness to the second. Now let R_1 be the relation:

$$[\lambda rs\,(r \neq u \,\&\, s \neq v \,\&\, Rrs) \vee (r = a \,\&\, s = b) \vee (r = u \,\&\, s = v)]$$

By our conventions for restricted variables, this λ-expression has the form $[\lambda xy\, D!x \,\&\, D!y \,\&\, \varphi]$, and so R_1 exists by (273.15). So it remains to show that R_1 is a witness to our **Claim** by showing both (A) and (B) hold with respect to R_1.

However, we can simplify our task by showing both R_1uv and that R_1 correlates$_D$ one-to-one F and G. For if so, then R_1 serves as the relevant witness and we have thereby reduced this problem to the reasoning in **Case 1**.

Clearly, R_1uv, by the third disjunct of R_1. So to show that R_1 correlates$_D$ one-to-one F and G, it suffices to note that R_1 is identical to R except that it maps a uniquely to b (by the first and second disjuncts of R_1) and u uniquely to v (by the first and third disjuncts of R_1). Hence, by the reasoning in **Case 1**, R_1 correlates$_D$ one-to-one F^{-u} and G^{-v}. ⋈

(754) Since the following proof is conducted entirely using variables and constants, all of which are significant, we may omit any existence clauses that would otherwise be required by the definitions of the notions involved. We continue to use r, s, and t as additional restricted variables for discernible objects, so that r, s, t, u, v all range over discernible objects. Assume that $F^{-u} \approx_D G^{-v}$, Fu, and Gv. The first implies, by definition (747.3) that some relation, say R, correlates$_D$ F^{-u} and G^{-v} one-to-one. Then by definition (747.2), we know:

Fact 1: $\forall r(F^{-u}r \rightarrow \exists!s(G^{-v}s \,\&\, Rrs)) \,\&\, \forall s(G^{-v}s \rightarrow \exists!r(F^{-u}r \,\&\, Rrs))$

Now by (750), it also follows that R maps$_D$ F^{-u} onto G^{-v} one-to-one. By definition (749.2), the one-to-one character of R entails:

Fact 2: $\forall r\forall s\forall t((F^{-u}r \,\&\, F^{-u}s \,\&\, G^{-v}t) \rightarrow (Rrt \,\&\, Rst \rightarrow r = s))$

We want to show $F \approx_D G$, i.e., by definitions (747.3) and (747.2):

Claim: There is a relation R' such that:

(A) $\forall r(Fr \rightarrow \exists!s(Gs \,\&\, R'rs))$

(B) $\forall s(Gs \rightarrow \exists!r(Fr \,\&\, R'rs))$

Consider the following relation R_2:

$$[\lambda rs\,(F^{-u}r \,\&\, G^{-v}s \,\&\, Rrs) \vee (r = u \,\&\, s = v)] \qquad R_2$$

By our convention for restricted variables, R_2 has the form $[\lambda xy\, D!x \,\&\, D!y \,\&\, \varphi]$, and so R_2 exists by (273.15). It remains to establish that R_2 is a witness to the **Claim**. But clearly, for any F-object a other than u, the first disjunct of R_2 guarantees that there is a unique G-object b other than v such that R_2ab, and vice versa. And by the second disjunct, R_2 uniquely maps u to v. Hence, R_2 correlates$_D$ F and G one-to-one.

(755.1) We prove our theorem by letting L be the property $[\lambda x\, Ex \rightarrow E!x]$ and showing that the following two properties are witnesses:

$\overline{L}$, i.e., $\overline{[\lambda x\, E!x \rightarrow E!x]}$

P: $[\lambda x\, E!x \,\&\, \neg\mathcal{A}E!x]$

Before we begin the proof proper, we note some facts about these two properties. Clearly $\overline{L}$ is an impossible property, i.e., $\Box\neg\exists x\overline{L}x$. Now as to P, recall that by theorem (211.1), we know $Contingent(q_0)$, where q_0 is $\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$. Hence by theorem (207.2), we know:

$\Diamond q_0 \,\&\, \Diamond\neg q_0$

i.e.,

(ω) $\Diamond\exists x(E!x \,\&\, \neg\mathcal{A}E!x) \,\&\, \Diamond\neg\exists x(E!x \,\&\, \neg\mathcal{A}E!x)$

Clearly, since $P\!\downarrow$, it follows by β-Conversion that $Px \equiv E!x \,\&\, \neg\mathcal{A}E!x$. From this modally strict fact it follows from (ω) by the Rule of Substitution (160.2) that:

(ϑ) $\Diamond\exists xPx \,\&\, \Diamond\neg\exists xPx$

Now to show that $\overline{L}$ and P are witnesses, we have to show $\Diamond(\overline{L} \approx_D P \,\&\, \Diamond\neg\overline{L} \approx_D P)$. But by (165.11), it suffices to show $\Diamond\overline{L} \approx_D P \,\&\, \Diamond\neg\overline{L} \approx_D P$:

- $\Diamond\overline{L} \approx_D P$. We first establish $\neg\exists xPx \rightarrow \overline{L} \approx_D P$. So assume $\neg\exists xPx$. *A fortiori*, $\neg\exists uPu$. Now we know, by definition of $\overline{L}$, that $\neg\exists x\overline{L}x$. Again, *a fortiori*, $\neg\exists u\overline{L}u$. So we may invoke (751.1) to conclude $P \approx_D \overline{L}$, which by the symmetry of $\approx_D$ (748.2), yields $\overline{L} \approx_D P$. So by Conditional Proof, $\neg\exists xPx \rightarrow \overline{L} \approx_D P$, and since this is a modally strict result, it follows by RM$\Diamond$ (157.2) that $\Diamond\neg\exists xPx \rightarrow \Diamond\overline{L} \approx_D P$. But we know $\Diamond\neg\exists xPx$, by (ϑ). Hence $\Diamond\overline{L} \approx_D P$.

- $\Diamond\neg\overline{L} \approx_D P$. We begin by first proving $\exists uPu \rightarrow \neg\overline{L} \approx_D P$. So assume $\exists uPu$. But we also know $\neg\exists x\overline{L}x$, which implies, *a fortiori*, that $\neg\exists u\overline{L}u$. So we may invoke (751.2) to conclude $\neg P \approx_D \overline{L}$, which by symmetry yields $\neg\overline{L} \approx_D P$. Hence, $\exists uPu \rightarrow \neg\overline{L} \approx_D P$. Since this is a modally strict result, it follows by RM$\Diamond$ (157.2) that implies $\Diamond\exists uPu \rightarrow \Diamond\neg\overline{L} \approx_D P$. But we know $\Diamond\exists xPx$, by (ϑ). But if $\Diamond\exists xPx$, then $\Diamond\exists uPu$ (exercise). Hence $\Diamond\neg\overline{L} \approx_D P$. ⋈

(755.2) We prove our theorem by again showing that the following two properties are witnesses:

$\overline{L}$: $\overline{[\lambda x\, E!x \rightarrow E!x]}$

P: $[\lambda x\, E!x \,\&\, \neg\mathcal{A}E!x]$

This time we have to show:

(ξ) $\Diamond([\lambda z\, \mathcal{A}\overline{L}z] \approx_D P \,\&\, \Diamond\neg[\lambda z\, \mathcal{A}\overline{L}z] \approx_D P)$

Note that the property $[\lambda z\, \mathcal{A}\overline{L}z]$ is, like $\overline{L}$, an impossible property, i.e., it is not hard to show $\Box\neg\exists x[\lambda z\, \mathcal{A}\overline{L}z]x$:[478]

> *Proof*. Suppose, for reductio, $[\lambda z\, \mathcal{A}\overline{L}z]x$. Since this property exists, it follows by β-Conversion that $\mathcal{A}\overline{L}x$. Hence, by (164.4), $\mathcal{A}\overline{L}x$. But, by definition of $\overline{L}$, we know $\Box\neg\overline{L}x$. Hence $\mathcal{A}\neg\overline{L}x$, by (132). So by axiom (44.1), $\neg\mathcal{A}\overline{L}x$. Contradiction. Hence $\neg[\lambda z\, \mathcal{A}\overline{L}z]x$, and by $\forall$I and quantificational logic, $\neg\exists x[\lambda z\, \mathcal{A}\overline{L}z]x$. Since this was established by modally strict means from no assumptions, RN implies $\Box\neg\exists x[\lambda z\, \mathcal{A}\overline{L}z]x$.

So $[\lambda z\, \mathcal{A}\overline{L}z]$ is an impossible property (i.e., necessarily unexemplified) whereas P is possibly exemplified and possibly not. Consequently, by reasoning analogous to that used in the two bullet points at the end of the proof of (755.1), we can show both conjuncts of (ξ). ⋈

(757.1) – **(757.2)** (Exercises)

(758.1)★ Since $F{\downarrow}$ and $[\lambda z\, \mathcal{A}Fz]{\downarrow}$, we have to show $\forall u([\lambda z\, \mathcal{A}Fz]u \equiv Fu)$, by definition (756). But this would follow, *a fortiori*, if we can show $\forall x([\lambda z \mathcal{A}Fz]x \equiv Fx)$. So, by GEN, we need only show $[\lambda z\, \mathcal{A}Fz]x \equiv Fx$. By β-Conversion and the fact that $[\lambda z\, \mathcal{A}Fz]{\downarrow}$, it follows that $[\lambda z\, \mathcal{A}Fz]x \equiv \mathcal{A}Fx$. But theorem (130.2)★ is $\mathcal{A}Fx \equiv Fx$. Hence, $[\lambda z\, \mathcal{A}Fz]x \equiv Fx$. ⋈

(758.2)★ (Exercise)

(759.1)[479] By theorem (133.4), we know $\mathcal{A}(\mathcal{A}Fx \equiv Fx)$, and since the commutativity of the biconditional is a modally strict theorem, it follows by a Rule of

[478] We thank Daniel West for suggesting a way to simplify the following subproof.

[479] In an earlier version, we had offered a proof of this theorem that appealed to the Fact discussed in Remark (137); the proof inferred $F \approx_D [\lambda z\, \mathcal{A}Fz]$ from (758.2)★, but then used the Fact just referenced to conclude $\mathcal{A}(F \approx_D [\lambda z \mathcal{A}Fz])$ is a modally strict theorem. But Daniel West reminded us (personal communication, 5 August 2022) that (a) there are reasonable extensions of object theory in which the Fact in question fails, as discussed in Remark (137), and (b) since many theorems of number theory rely on the present theorem, our proof risked undermining the application to number theory under reasonable extensions of object theory. So he presented an alternative proof that doesn't appeal to the Fact in Remark (137).

As I was about to input his proof and give him credit, I noticed that we had *commented out* our original proof of this theorem, which didn't appeal to the Fact in question. Further exmination showed that our original and Daniel's proof were almost identical – he saved one line at the outset, by appealing to (133.4) instead of (44.4), but we saved some lines in the middle of the proof by appealing to the Rule $\equiv$S of Biconditional Simplification! So the proof that follows is blend of the two. We'd like to thank Daniel for reminding us about the fragility of the Fact discussed in (137).

Substitution that $\mathcal{A}(Fx \equiv \mathcal{A}Fx)$. Since $[\lambda z\, \mathcal{A}Fz]{\downarrow}$, we know independently, by β-Conversion and the commutativity of the biconditional, that $\mathcal{A}Fx \equiv [\lambda z\, \mathcal{A}Fz]x$ is a modally strict theorem. So by a Rule of Substitution, we may infer $\mathcal{A}(Fx \equiv [\lambda z\mathcal{A}Fz]x)$. It follows by GEN that $\forall x\mathcal{A}(Fx \equiv [\lambda z\mathcal{A}Fz]x)$. So $\mathcal{A}\forall x(Fx \equiv [\lambda z\mathcal{A}Fz]x)$, by (44.3). *A fortiori*:

(ϑ) $\mathcal{A}\forall u(Fu \equiv [\lambda z\, \mathcal{A}Fz]u)$

Note that since $F{\downarrow}$ and $[\lambda z\, \mathcal{A}Fz]{\downarrow}$, definition (756) implies, by the Rule $\equiv$S of Biconditional Simplication (91), that $(F \equiv_D [\lambda z\, \mathcal{A}Fz]) \equiv \forall u(Fu \equiv [\lambda z\, \mathcal{A}Fz]u)$ is a modally strict theorem. Once we commute this result, it follows by a Rule of Substitution from (ϑ) that $\mathcal{A}(F \equiv_D [\lambda z\mathcal{A}Fz])$. But independently, if we apply the Rule of Actualization to (757.1), we know $\mathcal{A}(F \equiv_D [\lambda z\, \mathcal{A}Fz] \rightarrow F \approx_D [\lambda z\, \mathcal{A}Fz])$. Hence, by theorem (131), $\mathcal{A}(F \equiv_D [\lambda z\mathcal{A}Fz]) \rightarrow \mathcal{A}(F \approx_D [\lambda z\mathcal{A}Fz])$. But since we've established the antecedent, it follows that $\mathcal{A}(F \approx_D [\lambda z\, \mathcal{A}Fz])$. ⋈

(759.2) Assume $[\lambda z\, \mathcal{A}Fz]x$. Independently, since $[\lambda z\, \mathcal{A}Fz]{\downarrow}$, β-Conversion implies that the following as a modally strict theorem:

(ϑ) $[\lambda z\, \mathcal{A}Fz]x \equiv \mathcal{A}Fx$

So $\mathcal{A}Fx$. But this implies, by axiom (46.1), that $\Box\mathcal{A}Fx$. Since the commuted form of (ϑ) is also a modally strict theorem, it follows that $\Box[\lambda z\, \mathcal{A}Fz]x$, by a Rule of Substitution. So by conditional proof, $[\lambda z\, \mathcal{A}Fz]x \rightarrow \Box[\lambda z\, \mathcal{A}Fz]x$. Since this is a theorem, it follows by GEN, $\forall x([\lambda z\, \mathcal{A}Fz]x \rightarrow \Box[\lambda z\, \mathcal{A}Fz]x)$. And since this last result was proved by modally strict means, it follows by RN that $\Box\forall x([\lambda z\, \mathcal{A}Fz]x \rightarrow \Box[\lambda z\, \mathcal{A}Fz]x)$. So by definition (571.1), *Rigid*$([\lambda z\, \mathcal{A}Fz]$ ⋈

(760.1) Assume *Rigid*(F). Then it follows from definition (571.1) that $\Box\forall x(Fx \rightarrow \Box Fx)$. To show $F \approx_D [\lambda z\, \mathcal{A}Fz]$, it suffices by (757.1) to show $F \equiv_D [\lambda z\, \mathcal{A}Fz]$. Since this follows *a fortiori* from $\forall x(Fx \equiv [\lambda z\, \mathcal{A}Fz]x)$, we show the latter. Now our assumption implies, by CBF, that $\forall x\Box(Fx \rightarrow \Box Fx)$, and so by Rule $\forall$E, $\Box(Fx \rightarrow \Box Fx)$. But then by (174.2), it follows that $\mathcal{A}Fx \equiv Fx$. But we also know that since $[\lambda z\, \mathcal{A}Fz]{\downarrow}$, β-Conversion implies $[\lambda z\, \mathcal{A}Fz]x \equiv \mathcal{A}Fx$. Hence, by hypothetical syllogism, $[\lambda z\, \mathcal{A}Fz]x \equiv Fx$, which by symmetry yields $Fx \equiv [\lambda z\, \mathcal{A}Fz]x$. Since x isn't free in any assumption, it follows by GEN that $\forall x(Fx \equiv [\lambda z\mathcal{A}Fz]x)$. ⋈

(760.2) ($\rightarrow$) Assume $F \approx_D G$. Then by (748.4), we know:

(ϑ) $\forall H(H \approx_D F \equiv H \approx_D G)$

But by (39.2), $[\lambda z\, \mathcal{A}Hz]{\downarrow}$ and, moreover, this term is substitutable for H in the matrix of (ϑ). Hence:

$$[\lambda z\, \mathcal{A}Hz] \approx_D F \equiv [\lambda z\, \mathcal{A}Hz] \approx_D G$$

But then, by GEN, $\forall H([\lambda z\, \mathcal{A}Hz] \approx_D F \equiv [\lambda z\, \mathcal{A}Hz] \approx_D G)$.

($\leftarrow$) Assume $\forall H([\lambda z\, \mathcal{A}Hz] \approx_D F \equiv [\lambda z\, \mathcal{A}Hz] \approx_D G)$. Independently, by theorem (573), we know:

$$\exists H(\mathit{Rigidifies}(H,F))$$

Let P be such a property, so that we know $\mathit{Rigidifies}(P,F)$. It follows by (571.2) that:

(ϑ) $\mathit{Rigid}(P) \,\&\, \forall x(Px \equiv Fx)$

Now if we instantiate P into our assumption, we obtain:

(ζ) $[\lambda z\, \mathcal{A}Pz] \approx_D F \equiv [\lambda z\, \mathcal{A}Pz] \approx_D G)$

Our proof strategy is as follows:

(A) Show $[\lambda z\, \mathcal{A}Pz] \approx_D F$ and conclude, by symmetry, that $F \approx_D [\lambda z\, \mathcal{A}Pz]$.

(B) Conclude $[\lambda z\, \mathcal{A}Pz] \approx_D G$ from first part of (A) and (ζ).

(C) Conclude $F \approx_D G$, from the second part of (A), (B), and transitivity of $\approx_D$ (748.3).

Since (B), (C), and the second part of (A) are all straightforward, it remains only to show the first part of (A), i.e., $[\lambda z\, \mathcal{A}Pz] \approx_D F$. We do this by a direct hypothetical syllogism. From the first conjunct of (ϑ) and (760.1), it follows that $P \approx_D [\lambda z \mathcal{A}Pz]$, which implies $[\lambda z \mathcal{A}Pz] \approx_D P$, by symmetry. But from the second conjunct of (ϑ) it follows *a fortiori* that $P \equiv_D F$, and so $P \approx_D F$, by (757.1). Then, $[\lambda z\, \mathcal{A}Pz] \approx_D F$, by hypthetical syllogism. ⋈

(760.3) Let φ and ψ be as follows:

$$\varphi = \mathit{Rigid}(F) \,\&\, \mathit{Rigid}(G)$$

$$\psi = F \approx_D G \rightarrow \Box F \approx_D G$$

Then we have to show $\varphi \rightarrow \Box\psi$. Our proof strategy is as follows:

(i) Show that there is a modally strict proof of $\varphi \rightarrow \psi$.

(ii) Conclude by RM that $\Box\varphi \rightarrow \Box\psi$.

(iii) Show that $\varphi \rightarrow \Box\varphi$.

(iv) Conclude that $\varphi \rightarrow \Box\psi$ by hypothetical syllogism from (iii) and (ii).

Now steps (ii) and (iv) are trivial. Step (iii) is derivable as follows. Assume φ. Then if we expand both conjuncts by definition (571.1), the result is something of the form $\Box\chi_1 \,\&\, \Box\chi_2$. By the 4 and T schemas we know both that $\Box\chi_1 \equiv \Box\Box\chi_1$ and that $\Box\chi_2 \equiv \Box\Box\chi_2$. Since these are modally strict, it follows from φ by a Rule of Substitution that $\Box\Box\chi_1 \,\&\, \Box\Box\chi_2$. Hence $\Box(\Box\chi_1 \,\&\, \Box\chi_2)$, by the right-to-left direction of (158.3). But this is just $\Box\varphi$.

So it remains to show (i). Assume φ. Then by definition (571.1), we know both:

(A) $\Box\forall x(Fx \to \Box Fx)$, i.e., by (574.1)
$\forall x(\Diamond Fx \to \Box Fx)$

(B) $\Box\forall x(Gx \to \Box Gx)$, i.e., by (574.1)
$\forall x(\Diamond Gx \to \Box Gx)$

Now further assume $F \approx_D G$. Then, by (747.3) and (747.2):

$$\exists R(\forall u(Fu \to \exists!v(Gv \,\&\, Ruv)) \,\&\, \forall v(Gv \to \exists!u(Fu \,\&\, Ruv)))$$

Suppose R_1 is such a relation, so that we know:

(C) $\forall u(Fu \to \exists!v(Gv \,\&\, R_1uv)) \,\&\, \forall v(Gv \to \exists!u(Fu \,\&\, R_1uv))$

Now our goal is to show $\Box F \approx_D G$, and so by (747.3) and (747.2), we have to show:

$$\Box\exists R(\forall u(Fu \to \exists!v(Gv \,\&\, R_1uv)) \,\&\, \forall v(Gv \to \exists!u(Fu \,\&\, R_1uv)))$$

By the Buridan schema (168.1), it suffices to show:

$$\exists R\Box(\forall u(Fu \to \exists!v(Gv \,\&\, R_1uv)) \,\&\, \forall v(Gv \to \exists!u(Fu \,\&\, R_1uv)))$$

So we need to find a relation that is provably a witness to this claim. By theorem (573.3), we know that there is relation, say R_2, such that $\mathit{Rigidifies}(R_2,R_1)$, i.e., by definition (571.2):

$$\mathit{Rigid}(R_2) \,\&\, \forall x\forall y(R_2xy \equiv R_1xy)$$

i.e., by definition (571.1) and &E:

(D) $\Box\forall x\forall y(R_2xy \to \Box R_2xy)$, i.e., by (574.1)
$\forall x\forall y(\Diamond R_2xy \to \Box R_2xy)$

(E) $\forall x\forall y(R_2xy \equiv R_1xy)$

To show that R_2 is the needed witness, we have to show:

$$\Box(\forall u(Fu \to \exists!v(Gv \,\&\, R_2uv)) \,\&\, \forall v(Gv \to \exists!u(Fu \,\&\, R_2uv)))$$

Now by the right-to-left direction of (158.3), it suffices to show both $\Box\forall u(Fu \to \exists!v(Gv \,\&\, R_2uv))$ and $\Box\forall v(Gv \to \exists!u(Fu \,\&\, R_2uv))$. Without loss of generality, we show only the first. We leave it as an exercise to show that BF (167.1) holds for our restricted variable, i.e., we leave it as an exercise to show that it suffices to prove $\forall u\Box(Fu \to \exists!v(Gv \,\&\, R_2uv))$, and so by GEN, to show $\Box(Fu \to \exists!v(Gv \,\&\, R_2uv))$. Assume not, i.e., $\Diamond\neg(Fu \to \exists!v(Gv \,\&\, R_2uv))$, i.e., $\Diamond(Fu \,\&\, \neg\exists!v(Gv \,\&\, R_2uv))$. Then, by (162.3):

(F) $\Diamond Fu \,\&\, \Diamond\neg\exists!v(Gv \,\&\, R_2uv)$

Now our strategy is to use the first conjunct of (F) to show the negation of the second conjunct, yielding our contradiction. From the first conjunct of (F) and the second line of (A), it follows that $\Box Fu$. Hence, Fu and so it follows from the first conjunct of (C) that $\exists!v(Gv \,\&\, R_1uv)$. Suppose a is such a discernible object, so that by the definition of the uniqueness quantifier, we know:

(G) $Ga \,\&\, R_1ua \,\&\, \forall v'(Gv' \,\&\, R_1uv' \to v' = a)$

Recall that our strategy is to show the negation of the second conjunct of (F). So we have to show $\Box\exists!v(Gv \,\&\, R_2uv)$. That is, we have to show:

$$\Box\exists v(Gv \,\&\, R_2uv \,\&\, \forall v'(Gv' \,\&\, R_2uv' \to v' = v))$$

Now we leave it as an exercise that the Buridan formula $\exists\alpha\Box\varphi \to \Box\exists\alpha\varphi$ (168.1) holds for our rigid restricted variable v, i.e., that $\exists v\Box\varphi \to \Box\exists v\varphi$ is provable. Hence, it suffices to show:

$$\exists v\Box(Gv \,\&\, R_2uv \,\&\, \forall v'(Gv' \,\&\, R_2uv' \to v' = v))$$

We now show that a is a witness to this claim, which means we have to show:

$$\Box(Ga \,\&\, R_2ua \,\&\, \forall v'(Gv' \,\&\, R_2uv' \to v' = a))$$

By now familiar reasoning, it suffices to show:

(a) $\Box Ga$

(b) $\Box R_2ua$

(c) $\Box\forall v'(Gv' \,\&\, R_2uv' \to v' = a)$

But (a) follows from (B) and the first conjunct of (G). To show (b), note that (E) and the second conjunct of (G) yield R_2ua. But this result and (D) imply (b). Note that to show (c), it suffices to show $\forall v'\Box(Gv' \,\&\, R_2uv' \to v' = a)$, and by GEN, show $\Box(Gv' \,\&\, R_2uv' \to v' = a)$. For reductio, suppose not, i.e., $\Diamond\neg(Gv' \,\&\, R_2uv' \to v' = a)$, i.e., $\Diamond(Gv' \,\&\, R_2uv' \,\&\, \neg v' = a)$. Then:

(H) $\Diamond Gv' \,\&\, \Diamond R_2uv' \,\&\, \Diamond\neg v' = a$

But the first conjunct of (H) implies, by the second line of (B), that $\Box Gv'$ and, hence, Gv'. And the second conjunct of (H) implies, by the second line of (D), $\Box R_2uv'$, and so R_2uv'. From this latter, (E) implies R_1uv'. Since we've established Gv' and R_1uv', it follows from the third conjunct of (G) that $v' = a$. Then by the necessity of identity, $\Box v' = a$, i.e., $\neg\Diamond\neg v' = a$, which contradicts the third conjunct of (H). $\bowtie$

(763.1) – **(763.2)** (Exercises)

(764.1) Assume $G \approx_D H$. We need to show $Numbers(x, G) \equiv Numbers(x, H)$.

($\rightarrow$) Assume $Numbers(x, G)$. Then, by definition (762), we know $A!x$, $G\downarrow$, and:

(ϑ) $\forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)$

Since $H\downarrow$, to show $Numbers(x, H)$ it suffices by GEN to show $xF \equiv [\lambda z \mathcal{A}Fz] \approx_D H$. Now ($\vartheta$) implies $xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G$. Independently, our first assumption, $G \approx_D H$, implies, by (748.4), that $[\lambda z\, \mathcal{A}Fz] \approx_D G \equiv [\lambda z\, \mathcal{A}Fz] \approx_D H$. So, $xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D H$.

($\leftarrow$) By analogous reasoning. $\bowtie$

(764.2) Assume $Numbers(x, G)$ and $Numbers(x, H)$. Then by definition (762), these imply, respectively:

$$\forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)$$

$$\forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D H)$$

Hence, by (99.11) and (99.10), it follows that:

$$\forall F([\lambda z\, \mathcal{A}Fz] \approx_D G \equiv [\lambda z\, \mathcal{A}Fz] \approx_D H)$$

So by (760.2), $G \approx_D H$. $\bowtie$

(765.1) Assume both $Numbers(x, G)$ and $Numbers(y, H)$. ($\rightarrow$) Assume $x = y$. Then $Numbers(x, H)$. Hence by (764.2), $G \approx_D H$. ($\leftarrow$) Assume $G \approx_D H$. Then our global assumptions to imply, respectively:

(ϑ) $A!x \,\&\, G\downarrow \,\&\, \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)$

(ζ) $A!y \,\&\, H\downarrow \,\&\, \forall F(yF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D H)$

Since $A!x$ and $A!y$, it suffices by theorem (245.2) to show $\forall F(xF \equiv yF)$, and by GEN, that $xF \equiv yF$:

> ($\rightarrow$) Assume xF. Then the right conjunct of (ϑ) implies $[\lambda z\, \mathcal{A}Fz] \approx_D G$. But from this and our local assumption, it follows that $[\lambda z\, \mathcal{A}Fz] \approx_D H$, by the transitivity of $\approx_D$ (748.3) and the existence of $[\lambda z\, \mathcal{A}Fz]$. Hence, by the right conjunct of (ζ), it follows that yF.

(←) Assume yF. Then the right conjunct of (ζ) implies $[\lambda z\, \mathcal{A}Fz] \approx_D H$. Now since $\approx_D$ is symmetric (748.2), our local assumption implies that $H \approx_D G$. But then by transitivity of $\approx_D$, it follows that $[\lambda z\, \mathcal{A}Fz] \approx_D G$. So by the right conjunct of (ϑ), xF. ⋈

(765.2) (→) Assume $\exists x(Numbers(x,F) \,\&\, Numbers(x,G))$. Suppose a is such an object, so that we know $Numbers(a,F) \& Numbers(a,G)$. Then $F \approx_D G$, by (764.2).

(←) Assume $F \approx_D G$. Now by (763.1), we know $\exists x Numbers(x,F)$. Suppose b is such an object, so that we know $Numbers(b,F)$. Our assumption and this last fact imply, by (764.1), $Numbers(b,G)$. Hence, $\exists x(Numbers(x,F) \& Numbers(x,G))$. ⋈

(765.3) (Exercise)

(766) Assume $G \equiv_D H$. Then, by (757.1), $G \approx_D H$. Hence, our conclusion follows by (764.1). ⋈

(768) Assume $\exists u \exists v(u \neq v)$. Let c and d be such discernible objects, so that we know $c \neq d$. Then consider the properties $[\lambda x\, x = c]$ and $[\lambda x\, x = d]$, both of which exist (273.34). By Comprehension for Abstract Objects (53), we know that both:

$$\exists x(A!x \,\&\, \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D [\lambda x\, x = c]))$$

$$\exists x(A!x \,\&\, \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D [\lambda x\, x = d]))$$

Let a and b be such objects, so that we know, respectively:

$$A!a \,\&\, \forall F(aF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D [\lambda x\, x = c])$$

$$A!b \,\&\, \forall F(bF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D [\lambda x\, x = d])$$

Since we also know $[\lambda x\, x = c]\downarrow$ and $[\lambda x\, x = d]\downarrow$, it follows by definition (762), respectively, that:

(ϑ) $Numbers(a, [\lambda x\, x = c])$

(ξ) $Numbers(b, [\lambda x\, x = d])$

Our first goal is to show $a = b$. But by a Principle Underlying Hume's Principle (765.1), it now follows from (ϑ) and (ξ) that:

$$a = b \equiv [\lambda x\, x = c] \approx_D [\lambda x\, x = d]$$

But, clearly, $[\lambda x\, x = c] \approx_D [\lambda x\, x = d]$: the relation $[\lambda yz\, D!y \,\&\, D!z \,\&\, y = c \,\&\, z = d]$, which exists (273.15), correlates$_D$ the discernible objects exemplifying $[\lambda x x = c]$ and the discernible objects exemplifying $[\lambda x x = d]$ one-to-one (exercise). Hence $a = b$, and so it follows from (ξ) that:

(ζ) $Numbers(a,[\lambda x\, x = d])$

So, given (ϑ) and (ζ), it remains, by $\exists$I, to show $\neg[\lambda x\, x = c] \equiv_D [\lambda x\, x = d]$. So we have to show that some discernible object exemplifies one but not the other. But clearly, c exemplifies $[\lambda x x = c]$ but not $[\lambda x x = d]$, and d exemplifies $[\lambda x x = d]$ but not $[\lambda x\, x = c]$. ⋈

(769.1) By (755.2), we know $\exists F \exists G \Diamond([\lambda z\, \mathcal{A}Fz] \approx_D G \;\&\; \Diamond\neg[\lambda z\, \mathcal{A}Fz] \approx_D G)$. Suppose P and Q are such properties, so that we know:

$$\Diamond([\lambda z\, \mathcal{A}Pz] \approx_D Q \;\&\; \Diamond\neg[\lambda z\, \mathcal{A}Pz] \approx_D Q)$$

By (165.11), it follows that:

(ϑ) $\Diamond[\lambda z\, \mathcal{A}Pz] \approx_D Q$

(ζ) $\Diamond\neg[\lambda z\, \mathcal{A}Pz] \approx_D Q$

Now, by (763.1), we know $\exists x Numbers(x, Q)$. So suppose $Numbers(a, Q)$. By &I and two applications of $\exists$I, it remains to show $\neg\Box Numbers(a, Q)$. For reductio, assume $\Box Numbers(a, Q)$. Then by definition (762), $\Box(A!a \;\&\; Q\!\downarrow \;\&\; \forall F(aF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D Q))$. By (158.3), $\Box A!a$, $\Box Q\!\downarrow$, and $\Box\forall F(aF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D Q)$. By CBF (167.2), the latter implies $\forall F \Box(aF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D Q)$. Instantiating to P, we then know $\Box(aP \equiv [\lambda z\, \mathcal{A}Pz] \approx_D Q)$. By (158.4), this implies both:

(A) $\Box(aP \to [\lambda z\, \mathcal{A}Pz] \approx_D Q)$

(B) $\Box([\lambda z\, \mathcal{A}Pz] \approx_D Q \to aP)$

By the modally strict laws of contraposition and the Rule of Subsititution, (A) implies (C):

(C) $\Box(\neg[\lambda z\, \mathcal{A}Pz] \approx_D Q \to \neg aP)$

But, then (C) and (ζ) imply $\Diamond\neg aP$, by (158.13). So $\neg aP$, by (179.8). Analogously, (B) and (ϑ) imply $\Diamond aP$, by (158.13). So aP, by (179.3). Contradiction. ⋈

(769.2) Assume $Rigid(G)$, i.e., by definition (571.1), that $\Box\forall z(Gz \to \Box Gz)$ and assume $Numbers(x, G)$. It follows from the latter by definition that $A!x$, $G\!\downarrow$, and:

(ϑ) $\forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)$

Now we want to show $\Box Numbers(x, G)$, i.e.,

$$\Box(A!x \;\&\; G\!\downarrow \;\&\; \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G))$$

Since $\Box\varphi \;\&\; \Box\psi \;\&\; \Box\chi$ implies $\Box(\varphi \;\&\; \psi \;\&\; \chi)$, it suffices to show:

- $\Box A!x$

- $\Box G\downarrow$
- $\Box\forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)$

But the first follows from the fact that $A!x$ by (180.2) and the second is axiomatic (it is the modal closure of $G\downarrow$). So it remains to show the third. By BF, it suffices to show: $\forall F\Box(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)$. By GEN, it suffices to show $\Box(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)$. But we already know $\Box(xF \to \Box xF)$, by an application of RN to axiom (51). And note further that we can establish:

(ζ) $\Box([\lambda z\, \mathcal{A}Fz] \approx_D G \to \Box([\lambda z\, \mathcal{A}Fz] \approx_D G))$

Proof. We know $[\lambda z\, \mathcal{A}Fz]$ and G are both rigid properties: the first by (759.2) and the second by hypothesis. So by a relevant instance of (760.3), it follows that (ζ).

Then by an instance of theorem (172.5), where φ is xF and ψ is $[\lambda z\, \mathcal{A}Fz] \approx_D G$, it follows that $\Box((\varphi \equiv \psi) \to \Box(\varphi \equiv \psi))$, i.e.,

$$\Box((xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G) \to \Box(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)$$

Then by the T schema, it follows that:

$$(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G) \to \Box(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)$$

Since the antecedent follows from (ϑ), we have established:

$$\Box(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)$$

which was what we had to show. ⋈

(769.3) (Exercise)

(769.4) We begin with some preliminary results that prepare the ground for our reasoning. If we apply RN and then GEN (2x) to the modally strict theorem (765.3) and then instantiate the result to G and $[\lambda z\, \mathcal{A}Gz]$, we know:

$$\Box(\exists x(Numbers(x, G)\,\&\, Numbers(x, [\lambda z\, \mathcal{A}Gz])) \equiv G \approx_D [\lambda z\, \mathcal{A}Gz])$$

So by (132), it follows that:

(A) $\mathcal{A}(\exists x(Numbers(x, G)\,\&\, Numbers(x, [\lambda z\, \mathcal{A}Gz])) \equiv G \approx_D [\lambda z\, \mathcal{A}Gz])$

Now by (759.1), we know $\mathcal{A}(G \approx_D [\lambda z\, \mathcal{A}Gz])$. Hence from this fact and (A) it follows by the relevant instance of (139.5) that:

$$\mathcal{A}\exists x(Numbers(x, G)\,\&\, Numbers(x, [\lambda z\, \mathcal{A}Gz]))$$

From this it follows by (139.10) that:

$$\exists x\mathcal{A}(Numbers(x, G)\,\&\, Numbers(x, [\lambda z\, \mathcal{A}Gz]))$$

Suppose a is such an object, so that we know:

$$\mathcal{A}(Numbers(a,G)\,\&\,Numbers(a,[\lambda z\,\mathcal{A}Gz]))$$

Then by (139.2) and &E, it follows that:

(B) $\mathcal{A}Numbers(a,G)$

(C) $\mathcal{A}Numbers(a,[\lambda z\,\mathcal{A}Gz])$

Now, independently, by applying CBF (167.2) to (769.3) and instantiating the result to a, we know that $Numbers(a,[\lambda z\,\mathcal{A}Gz])$ is modally collapsed, i.e., that:

$$\Box(Numbers(a,[\lambda z\,\mathcal{A}Gz]) \rightarrow \Box Numbers(a,[\lambda z\,\mathcal{A}Gz]))$$

From this last result and (C) we may infer:

(D) $Numbers(a,[\lambda z\,\mathcal{A}Gz])$

With these results in hand, we may prove our theorem as follows: ($\rightarrow$) Suppose $\mathcal{A}Numbers(x,G)$. Now since (763.2) is a theorem, we may apply the Rule of Actualization to infer $\mathcal{A}\exists!yNumbers(y,G)$. So by (176.1), $\exists!y\mathcal{A}Numbers(y,G)$. Hence, our assumption and (B) imply $a=x$. So by (D), $Numbers(x,[\lambda z\,\mathcal{A}Gz])$. ($\leftarrow$) Suppose $Numbers(x,[\lambda z\mathcal{A}Gz])$. But (763.2) ensures $\exists!yNumbers(y,[\lambda z\mathcal{A}Gz])$. So our assumption and (D) imply $a=x$. So by (B), $\mathcal{A}Numbers(x,G)$. ⋈

(770) (Exercise)

(771.2) (Exercise)

(772.1) By (770), we know:

(ϑ) $\imath xNumbers(x,G)\downarrow$

Independently, given that $G\downarrow$ is axiomatic, we know that definition (762) implies the following, simplified biconditional:

$$Numbers(x,G) \equiv (A!x\,\&\,\forall F(xF \equiv [\lambda z\,\mathcal{A}Fz] \approx_D G)))$$

Since this is a modally strict theorem, it follows by GEN and RN that:

(ξ) $\Box\forall x(Numbers(x,G) \equiv A!x\,\&\,\forall F(xF \equiv [\lambda z\,\mathcal{A}Fz] \approx_D G))$

So from (ϑ) and (ξ), it follows by (149.3) that:

(ζ) $\imath xNumbers(x,G) = \imath x(A!x\,\&\,\forall F(xF \equiv [\lambda z\,\mathcal{A}Fz] \approx_D G))$

Since all of the terms in question are significant, it follows by definition (771.1), (ζ), and the transitivity of identity that:

$$\#G = \imath x(A!x\,\&\,\forall F(xF \equiv [\lambda z\,\mathcal{A}Fz] \approx_D G))$$ ⋈

(772.2) We just established:

$$\#G = \iota x(A!x \,\&\, \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G))$$

If we can establish:

$$\iota x(A!x \,\&\, \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)) = \iota x(A!x \,\&\, \forall F(xF \equiv F \approx_D G))$$

then by transitivity of identity, we're done. Our proof strategy to establish the above identity is to use an appropriate instance of (149.2). We already know:

$$\iota x(A!x \,\&\, \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G))\downarrow$$

since this follows from the preceding theorem. So it remains to show the following equivalence:

$$\mathcal{A}\forall x((A!x \,\&\, \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)) \equiv (A!x \,\&\, \forall F(xF \equiv F \approx_D G)))$$

But the actuality operator commutes with the universal quantifier, so it suffices to show:

$$\forall x\mathcal{A}((A!x \,\&\, \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)) \equiv (A!x \,\&\, \forall F(xF \equiv F \approx_D G)))$$

Then by GEN, it suffices to show:

$$\mathcal{A}((A!x \,\&\, \forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)) \equiv (A!x \,\&\, \forall F(xF \equiv F \approx_D G)))$$

So by (139.5) and (139.2), it suffices to show:

$$(\mathcal{A}A!x \,\&\, \mathcal{A}\forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G)) \equiv (\mathcal{A}A!x \,\&\, \mathcal{A}\forall F(xF \equiv F \approx_D G))$$

Simplifying (i.e., removing $A!x$ from both sides), we want to show:

$$\mathcal{A}\forall F(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G) \equiv \mathcal{A}\forall F(xF \equiv F \approx_D G)$$

Again, by (44.3) and GEN, it suffices to show:

$$\mathcal{A}(xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D G) \equiv \mathcal{A}(xF \equiv F \approx_D G)$$

Again, by (139.5), we have to show:

$$(\mathcal{A}xF \equiv \mathcal{A}[\lambda z\, \mathcal{A}Fz] \approx_D G) \equiv (\mathcal{A}xF \equiv \mathcal{A}F \approx_D G)$$

Since it is a theorem of propositional logic that $((\varphi \equiv \psi) \equiv (\varphi \equiv \chi)) \equiv (\psi \equiv \chi)$, we simply need to show:

$$(\mathcal{A}[\lambda z\, \mathcal{A}Fz] \approx_D G) \equiv (\mathcal{A}F \approx_D G)$$

Again, by (139.5), we have to show:

$$(\vartheta)\quad \mathcal{A}([\lambda z\, \mathcal{A}Fz] \approx_D G \equiv F \approx_D G)$$

Now, consider the following consequence of (760.2):

$$\forall H([\lambda z\, \mathcal{A}Fz] \approx_D F \equiv (H \approx_D [\lambda z\, \mathcal{A}Fz] \equiv H \approx_D F))$$

Instantiating to G, we have $([\lambda z\, \mathcal{A}Fz] \approx_D F) \equiv (G \approx_D [\lambda z\, \mathcal{A}Fz] \equiv G \approx_D F)$. By the commutativity of $\equiv$:

$$(G \approx_D [\lambda z\, \mathcal{A}Fz] \equiv G \approx_D F) \equiv ([\lambda z\, \mathcal{A}Fz] \approx_D F)$$

Since it is a modally strict theorem that $\approx_D$ is a symmetrical condition (273.31), it follows by a Rule of Substitution that:

$$([\lambda z\, \mathcal{A}Fz] \approx_D G \equiv F \approx_D G) \equiv (F \approx_D [\lambda z\, \mathcal{A}Fz])$$

Since we proved the above as a theorem from no premises, it follows by the Rule of Actualiation (RA) and the distribution of $\mathcal{A}$ over a biconditional (139.5) that:

$$\mathcal{A}([\lambda z\, \mathcal{A}Fz] \approx_D G \equiv F \approx_D G) \equiv \mathcal{A}(F \approx_D [\lambda z\, \mathcal{A}Fz])$$

Given this fact, to show (ϑ) we need only show:

$$\mathcal{A}(F \approx_D [\lambda z\, \mathcal{A}Fz])$$

But this is just (759.1).

(774.1) Since y is a variable substitutable for x in $Numbers(x,G)$ and doesn't occur free in $Numbers(x,G)$, we have the following alphabetic variant of an instance of axiom (47), where φ in that theorem is set to $Numbers(x,G)$:

$$x = \imath x Numbers(x,G) \ \equiv\ \forall y(\mathcal{A}Numbers(y,G) \equiv y = x)$$

But by (769.4), $\mathcal{A}Numbers(y,G) \equiv Numbers(y,[\lambda z\, \mathcal{A}Gz])$ is a modally strict theorem. Hence by a Rule of Substitution, it follows from our first displayed line that:

$$x = \imath x Numbers(x,G) \ \equiv\ \forall y(Numbers(y,[\lambda z\, \mathcal{A}Gz]) \equiv y = x)$$

Since this is a theorem, it follows by GEN that:

$$(\vartheta)\quad \forall x(x = \imath x Numbers(x,G) \ \equiv\ \forall y(Numbers(y,[\lambda z\, \mathcal{A}Gz]) \equiv y = x))$$

But by (770), $\imath x Numbers(x,G)\downarrow$. Hence by definition (771.1) and Rule =E, we know $\#G\downarrow$. From this last result and (ϑ), it follows by Rule $\forall$E (93.1) Variant that:

$$\#G = \imath x Numbers(x,G) \ \equiv\ \forall y(Numbers(y,[\lambda z\, \mathcal{A}Gz]) \equiv y = \#G)$$

So, by definition (771.1) and biconditional syllogism, it follows that:

$$\forall y(Numbers(y,[\lambda z\, \mathcal{A}Gz]) \equiv y = \#G)$$

By instantiating this last result to x, and we have:

$$Numbers(x,[\lambda z\, \mathcal{A}Gz]) \equiv x = \#G \qquad \bowtie$$

(774.2) – **(774.4)** (Exercises)

(774.5) Assume $Rigid(G)$. Then $\Box\forall x(Numbers(x,G) \to \Box Numbers(x,G))$, by theorem (769.2). So by the T schema:

$$(\vartheta)\quad \forall x(Numbers(x,G) \to \Box Numbers(x,G))$$

Then by (153.2):

$$\exists! x Numbers(x,G) \to (\forall y(y = \imath x Numbers(x,G) \to Numbers(y,G)))$$

But it is a theorem that $\exists! x Numbers(x,G)$ (763.2). Hence:

$$\forall y(y = \imath x Numbers(x,G) \to Numbers(y,G))$$

Since $\#G\downarrow$, we may instantiate it to obtain:

$$\#G = \imath x Numbers(x,G) \to Numbers(\#G,G))$$

Since the antecedent holds by definition (771.1), $Numbers(\#G,G)$. ⋈

(775) Assume both $Rigid(F)$ and $Rigid(G)$. Note that, independently, it follows from (765.1), by several applications of GEN, that:

$$\forall G\forall H\forall x\forall y[(Numbers(x,G) \,\&\, Numbers(y,H)) \to (x = y \equiv G \approx_D H)]$$

If we instantiate $\forall G$ to F, $\forall H$ to G, $\forall x$ to $\#F$, and $\forall y$ to $\#G$, we obtain:

$$(Numbers(\#F,F) \,\&\, Numbers(\#G,G)) \to (\#F = \#G \equiv F \approx_D G)$$

But (774.5) and our assumptions imply $Numbers(\#F,F)$ and $Numbers(\#G,G)$. So $\#F = \#G \equiv F \approx_D G$. ⋈

(776.1)★ By definition (771.1) and theorem (145.2)★. ⋈

(776.2)★ It follows from (765.1) by several applications of GEN that:

$$\forall G\forall H\forall x\forall y[(Numbers(x,G) \,\&\, Numbers(y,H)) \to (x = y \equiv G \approx_D H)]$$

If we instantiate $\forall G$ to F, $\forall H$ to G, $\forall x$ to $\#F$, and $\forall y$ to $\#G$, we obtain:

$$(Numbers(\#F,F) \,\&\, Numbers(\#G,G)) \to (\#F = \#G \equiv F \approx_D G)$$

By (776.1)★, we know both $Numbers(\#F, F)$ and $Numbers(\#G, G)$. So by &I and MP, $\#F = \#G \equiv F \approx_D G$. ⋈

(776.3)★ (Exercise)

(776.4)★ Assume $F \equiv_D G$. Then $F \approx_D G$, by (757.1). So $\#F = \#G$, by (776.2)★. ⋈

(778.1) Assume $Numbers(x, G)$. Independently, by theorem (573.3), we know $\exists F(Rigidifies(F, G))$. Let P be such a property, so that we know $Rigidifies(P, G)$. Then, by definition (571.2), we know:

(ϑ) $Rigid(P) \,\&\, \forall x(Px \equiv Gx)$

Since the second conjunct of (ϑ) implies that P and G are materially equivalent$_D$ (i.e., implies $P \equiv_D G$), $Numbers(x, P)$ follows from our initial assumption, by (766). Independently, the first conjunct of (ϑ) and theorem (760.1) imply that $P \approx_D [\lambda z\, \mathcal{A}Pz]$. It follows that $Numbers(x, [\lambda z\, \mathcal{A}Pz])$, by (764.1). So, by (774.1) we have $x = \#P$. Hence, $\exists F(x = \#F)$, and so $NaturalCardinal(x)$, by definition (777). ⋈

(778.2) ($\rightarrow$) Assume $\exists G(x = \#G)$. So suppose P is such a property, so that we know $x = \#P$. Then by (774.2), $Numbers(x, [\lambda z\, \mathcal{A}Pz])$. So $\exists G(Numbers(x, G))$. ($\leftarrow$) Assume $\exists G(Numbers(x, G))$, and let Q be such a property, so that we know $Numbers(x, Q)$. Then by (778.1), $NaturalCardinal(x)$, and so by definition (777), $\exists G(x = \#G)$. ⋈

(779.1) Assume $NaturalCardinal(x)$. Then by definition (777), $\exists G(x = \#G)$. Suppose P is such a property, so that we know $x = \#P$. It follows by the necessity of identity (125.2) that $\Box(x = \#P)$. Hence, by $\exists$I, $\exists G\Box(x = \#G)$. But then by the Buridan formula (168.1), $\Box\exists G(x = \#G)$. Hence by definition (777), $\Box NaturalCardinal(x)$. ⋈

(779.2) (Exercise)

(780) Assume $NaturalCardinal(x)$. So $\exists G(x = \#G)$, by (777). Let P be such a property, so that we know $x = \#P$. Hence, by (774.1):

(ϑ) $Numbers(x, [\lambda z\, \mathcal{A}Pz])$

By GEN, we have to show $xF \equiv x = \#F$. If we substitute P for G in (774.3), then we independently know:

$$\#PF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D [\lambda z\, \mathcal{A}Pz]$$

Hence:

(ξ) $xF \equiv [\lambda z\, \mathcal{A}Fz] \approx_D [\lambda z\, \mathcal{A}Pz]$

Next, we leave it as an exercise to show that as a matter of propositional logic, the following is a theorem:

(ζ) $\vartheta \to ((xF \equiv (\psi \equiv \vartheta)) \equiv (xF \equiv \psi))$

Hence, we may reason as follows:

$$\begin{array}{llll} xF & \equiv & [\lambda z\, \mathcal{A}Fz] \approx_D [\lambda z\, \mathcal{A}Pz] & \text{by } (\xi) \\ & \equiv & Numbers(x,[\lambda z\, \mathcal{A}Fz]) \equiv Numbers(x,[\lambda z\, \mathcal{A}Pz]) & \text{by } (764.1) \\ & \equiv & Numbers(x,[\lambda z\, \mathcal{A}Fz]) & \text{via } (\vartheta) \text{ and } (\zeta) \\ & \equiv & x = \#F & \text{by } (774.1) \end{array}$$

(781) (Exercise)

(783) By definition (782.1), $0 = \#[\lambda u\, u \neq_D u]$. Since $[\lambda u\, u \neq_D u]\downarrow$, it follows that $\exists G(0 = \#G)$. Hence, by definition (777), *NaturalCardinal*(0). ⋈

(784.1) ($\to$) Assume $\neg\exists uFu$. Note that $\neg\exists v([\lambda u\, \mathcal{A}u \neq_D u]v)$ is an easily established theorem:

> Clearly, since $v =_D v$ is a modally-strict theorem, then so is $\neg(v \neq_D v)$. Note that since $D!x \to \Box D!x$ (273.8), the variable v ranges over objects satisfying a rigid condition. Hence we can apply expanded RN (341.3.a) and conclude $\Box\neg(v \neq_D v)$ and by (132), $\mathcal{A}\neg(v \neq_D v)$ and hence by (44.1), $\neg\mathcal{A}(v \neq_D v)$. But then by Rule $\overleftarrow{\beta}$C, $\neg[\lambda u\, \mathcal{A}u \neq_D u]v$. So $\forall v \neg[\lambda u\, \mathcal{A}u \neq_D u]v$, by GEN, i.e., $\neg\exists v[\lambda u\, \mathcal{A}u \neq_D u]v$.

From this and our hypothesis that $\neg\exists uFu$, it follows that $[\lambda u\, \mathcal{A}u \neq_D u] \approx_D F$ by (751.1). By (764.1), it follows that $Numbers(0,[\lambda u\, \mathcal{A}u \neq_D u]) \equiv Numbers(0,F)$. So it remains only to show: $Numbers(0,[\lambda u\, \mathcal{A}u \neq_D u])$. But $0 = \#[\lambda u\, u \neq_D u]$, by definition. So by (774.2), $Numbers(0,[\lambda y\, \mathcal{A}[\lambda u\, u \neq_D u]y])$. But we now leave it as an exercise to show that $Numbers(0,[\lambda u\, \mathcal{A}u \neq_D u])$. (In completing the exercise you might show that if $Numbers(0,[\lambda y\, \varphi])$, then $Numbers(0,[\lambda u\, \varphi])$.)

($\leftarrow$) Assume $Numbers(0,F)$. By reasoning given above, we already know that $Numbers(0,[\lambda u\, \mathcal{A}u \neq_D u])$. So, by (764.2), we have $F \approx_D [\lambda u\, \mathcal{A}u \neq_D u]$. We also know, by reasoning given above, that $\neg\exists v[\lambda u\, \mathcal{A}u \neq_D u]v$. Now, suppose for reductio, that $\exists uFu$. Then, by (751.2), it follows that $F \not\approx_D [\lambda u\, \mathcal{A}u \neq_D u]$. Contradiction. ⋈

(784.2) ($\to$) Assume $\exists uFu$. Then by (763.1), $\exists x Numbers(x,F)$. Say it is b, so that $Numbers(b,F)$. By &I and $\exists$I, it remains to show $b \neq 0$. For reductio, suppose $b = 0$. Then $Numbers(0,F)$, by Rule =E. So by (784.1), $\neg\exists uFu$. Contradiction. ($\leftarrow$) Assume $\exists x(Numbers(x,F) \,\&\, x \neq 0)$. Suppose a is such, so that we know $Numbers(a,F) \,\&\, a \neq 0$. Now suppose, for reductio, $\neg\exists uFu$. Then by (784.1), $Numbers(0,F)$. But by (763.2), $\exists! x Numbers(x,F)$. So $a = 0$. Contradiction. ⋈

(784.3) – **(784.4)** (Exercises)

(784.5) ($\to$) Assume $w \models \neg\exists uFu$. Then $\neg w \models \exists uFu$, by theorem (529.1). But the modally strict theorem (545.6), i.e., $(w \models \exists xFx) \equiv \exists x(w \models Fx)$, implies that

$(w \models \exists uFu) \equiv \exists u(w \models Fu)$ is a modally strict theorem (exercise).[480] So by a Rule of Substitution, $\neg\exists u(w \models Fu)$. Moreover, a universally generalized (573.1) implies that $w \models Fu \equiv F_w u$ is also a modally strict theorem. So by a Rule of Substitution, $\neg\exists uF_w u$. But this implies $\Box\neg\exists uF_w u$. To see why, suppose not. Then $\Diamond\exists uF_w u$ and so by CBF, $\exists u\Diamond F_w u$ (recall that the Barcan Formulas hold for rigid restricted variables). Suppose a is such a discernible object so that we know $\Diamond F_w a$. But it is a modally strict theorem (573.2) that F_w is rigid and so the definition of rigidity (571.1) lets us conclude that $\Box\forall x(F_w x \to \Box F_w x)$. By CBF, $\forall x\Box(F_w x \to \Box F_w x)$, and so $\Box(F_w a \to \Box F_w a)$. By a fact about modal collapse, namely (172.1), it follows from this last fact and $\Diamond F_w a$ that $\Box F_w a$. By the T schema, $F_w a$ and, hence, $\exists uF_w u$, which contradicts $\neg\exists uF_w u$. So by reductio, $\Box\neg\exists uF_w u$. And this implies, by (784.4), that $\#(F_w) = 0$.

($\leftarrow$) Assume $\#(F_w) = 0$. So by (784.3), $\neg\exists u\mathcal{A}F_w u$. But we saw above that it is a modally strict theorem that F_w is rigid, and so $\Box\forall x(F_w x \to \Box F_w x)$. By CBF, $\forall x\Box(F_w x \to \Box F_w x)$. Hence $\Box(F_w u \to \Box F_w u)$. But from this and another fact about modal collapse (174.2), we can derive, as a modally strict theorem, that $\mathcal{A}F_w u \equiv F_w u$. So by a Rule of Substitution, $\neg\exists uF_w u$. Now by theorem (573.1), it follows *a fortiori* that $F_w u \equiv w \models Fu$ is also a modally strict theorem. Hence by a Rule of Substitution, $\neg\exists u(w \models Fu)$. So by (545.6) and a Rule of Substitution, $\neg(w \models \exists uFu)$. And by (529.1), $w \models \neg\exists uFu$. ⋈

(785.1)★ Assume $NaturalCardinal(x)$. It suffices to show $xF \equiv Numbers(x,F)$, by GEN. By (780), it follows from our assumption that:

$$\forall F(xF \equiv x = \#F)$$

So in particular, $xF \equiv x = \#F$. But by definition of $\#F$, this implies:

$$(\vartheta)\quad xF \equiv x = \imath y(Numbers(y,F))$$

But by (145.2)★, we know:

$$(x = \imath yNumbers(y,F)) \to Numbers(x,F)$$

Moreover, we leave it as an exercise, using (763.2) and the logic of descriptions, to show that $Numbers(x,F) \to x = \imath yNumbers(y,F)$. Hence:

[480] Here is a solution. We want to show $(w \models \exists uFu) \equiv \exists u(w \models Fu)$. ($\to$) Assume $w \models \exists uFu$, i.e., $w \models \exists x(D!x \,\&\, Fx)$. To show $\exists u(w \models Fu)$, we have to show $\exists x(D!x \,\&\, w \models Fx)$. Our assumption implies $\exists x(w \models (D!x \,\&\, Fx))$, by (545.6). Let a be such an object, so that we know $w \models (D!a \,\&\, Fa)$. Then by (545.1), we know both $w \models D!a$ and $w \models Fa$. The former implies $\exists w(w \models D!a)$, and so by a Fundamental Theorem, $\Diamond D!a$. Hence $D!a$, by (273.10). Since we now have $D!a \,\&\, w \models Fa$, it follows that $\exists x(D!x \,\&\, w \models Fx)$.

($\leftarrow$) Assume $\exists u(w \models Fu)$, i.e., $\exists x(D!x \& w \models Fx)$, to show $w \models \exists x(D!x \& Fx)$. Let b be such an object, so that we know both $D!b$ and $w \models Fb$. But the first implies $\Box D!b$, by (273.9). So by a Fundamental Theorem, $\forall w'(w' \models D!b)$. Hence $w \models D!b$. Since we've established $(w \models D!b) \,\&\, (w \models Fb)$, it follows by (545.1) that $w \models (D!b \,\&\, Fb)$. Hence, $\exists x(w \models (D!x \,\&\, Fx))$. And so by (545.6), $w \models \exists x(D!x \,\&\, Fx)$. ⋈

(ξ) $(x = \iota y Numbers(y,F)) \equiv Numbers(x,F)$

So from (ϑ) and (ξ), $xF \equiv Numbers(x,F)$. ⋈

(785.2)★ We know $NaturalCardinal(0)$. So, by (785.1)★, $0F \equiv Numbers(0,F)$. But, by (784.1), it follows that $0F \equiv \neg\exists u Fu$. ⋈

(785.3)★ By (785.2)★ and the symmetry of $\equiv$, we know that $\neg\exists u Fu \equiv 0F$. But by (780), $0F \equiv 0 = \#F$. It follows that $\neg\exists u Fu \equiv \#F = 0$. ⋈

(786.2) This is an instance of axiom (39.2), since the following is a core λ-expression:

$$[\lambda xy\ \forall F((\forall z(Gxz \to Fz)\ \&\ Hereditary(F,G)) \to Fy)]$$

To confirm this, note that no variable bound by the λ occurs in encoding position (9.1) in the matrix. Though the defined expression $Hereditary(F,G)$ does, by convention, contain encoding formulas (buried in the conjuncts $G\downarrow$ and $F\downarrow$ of its definiens), x and y are not free in those formulas. To confirm this, $Hereditary(F,G)$ is defined as (786.1):

$$F\downarrow\ \&\ G\downarrow\ \&\ \forall x\forall y(Gxy \to (Fx \to Fy))$$

where $F\downarrow$ and $G\downarrow$ are defined as (20.2), respectively:

$$\exists x(xF)$$

$$\exists x\exists y(xyG)$$

Though x and y occur in encoding position in xyG and xG, both are already bound and so not bound by the λ in the target λ-expression. ⋈

(788) (Exercise)

(789.1) Assume Gxy. By (788) and GEN, we have to show:

$$(\forall z(Gxz \to Fz)\ \&\ Hereditary(F,G)) \to Fy$$

So assume $\forall z(Gxz \to Fz)$ and $Hereditary(F,G)$. Instantiate the first of these to y, to obtain $Gxy \to Fy$. But then by our first assumption, Fy. ⋈

(789.2) Assume G^*xy, $\forall z(Gxz \to Fz)$, and $Hereditary(F,G)$. Then by the first assumption and the fundamental fact about G^* (788), we know:

$$\forall F[(\forall z(Gxz \to Fz)\ \&\ Hereditary(F,G)) \to Fy]$$

But we may instantiate this to F to obtain:

$$(\forall z(Gxz \to Fz)\ \&\ Hereditary(F,G)) \to Fy$$

But both conjuncts of the antecedent hold by assumption. ⋈

(789.3) Assume Fx, G^*xy, and that $Hereditary(F,G)$. Then by (789.2), it suffices to show only $\forall z(Gxz \rightarrow Fz)$. So assume Gxz. Since F is G-hereditary, we know by definition (786.1) that $\forall x \forall y(Gxy \rightarrow (Fx \rightarrow Fy))$. In particular, $Gxz \rightarrow (Fx \rightarrow Fz)$. But both Gxz and Fx both hold by assumption. ⋈

(789.4) Assume Gxy and G^*yz. To prove G^*xz, further assume $\forall z(Gxz \rightarrow Fz)$ and $Hereditary(F,G)$. Our first and third assumptions imply Fy. But from this, G^*yz, and $Hereditary(F,G)$, it follows that Fz, by (789.3). ⋈

(789.5) Assume G^*xy. Now, by (789.2) and GEN, we know:

$$\forall F[(G^*xy \,\&\, \forall z(Gxz \rightarrow Fz) \,\&\, Hereditary(F,G)) \rightarrow Fy]$$

Instantiate to $[\lambda y' \, \exists x' Gx'y']$ (which we know exists) and we obtain:

$$(G^*xy \,\&\, \forall z(Gxz \rightarrow [\lambda y' \, \exists x' Gx'y']z) \,\&\, Hereditary([\lambda y' \, \exists x' Gx'y'], G)) \rightarrow [\lambda y' \, \exists x' Gx'y']y$$

Since $[\lambda y' \, \exists x' Gx'y']$ exists, β-Conversion (181) and the Rule of Substitution (160.2) allow us to reduce this to:

$$(\vartheta) \quad (G^*xy \,\&\, \forall z(Gxz \rightarrow \exists x' Gx'z) \,\&\, Hereditary([\lambda y' \, \exists x' Gx'y'], G)) \rightarrow \exists x' Gx'y$$

The consequent of (ϑ), $\exists x' Gx'y$, is an alphabetic variant of what we have to show. So since G^*xy by assumption, it remains to show the second and third conjuncts of (ϑ). But these are quickly obtained. For the second conjunct, assume Gxz, by GEN. Then $\exists x' Gx'z$, by $\exists$I. For the third conjunct, we have to show, by (786.1):

$$\forall z \forall z'(Gzz' \rightarrow ([\lambda y' \, \exists x' Gx'y']z \rightarrow [\lambda y' \, \exists x' Gx'y']z'))$$

Since $[\lambda y' \, \exists x' Gx'y']$ exists, it suffices to show, by β-Conversion and the Rule of Substitution (160.2):

$$\forall z \forall z'(Gzz' \rightarrow (\exists x' Gx'z \rightarrow \exists x' Gx'z'))$$

By GEN, it suffices to show $Gzz' \rightarrow (\exists x' Gx'z \rightarrow \exists x' Gx'z')$. So assume Gzz' and $\exists x' Gx'z$. But the first of these assumptions yields $\exists x' Gx'z'$. ⋈

(789.6) Assume G^*xy and G^*yz. Then by theorem (788), it follows that:

$$(\vartheta) \quad \forall F[(\forall z(Gxz \rightarrow Fz) \,\&\, Hereditary(F,G)) \rightarrow Fy]$$

$$(\xi) \quad \forall F[(\forall z'(Gyz' \rightarrow Fz') \,\&\, Hereditary(F,G)) \rightarrow Fz]$$

To show G^*xz, we have to show, by GEN:

$$(\forall z'(Gxz' \rightarrow Fz') \,\&\, Hereditary(F,G)) \rightarrow Fz$$

So assume:

(ζ) $\forall z'(Gxz' \rightarrow Fz')$ & $Hereditary(F,G)$

to show Fz. Now with the second conjunct of (ζ) in hand, we need only to show $\forall z'(Gyz' \rightarrow Fz')$, for then by ($\xi$), we can infer Fz. So, by GEN, assume Gyz', to show Fz'. Now, independently, (ϑ) and (ζ) jointly imply Fy. But from Gyz' and Fy, it follows by the definition of $Hereditary(F,G)$ (786.1) that Fz'. ⋈

(789.7) We have to show $\underline{G}^*xy \rightarrow \exists z\underline{G}xz$. To minimize clash of variables, we prove $\underline{G}^*ab \rightarrow \exists z\underline{G}az$, where a and b are any arbitrarily chosen objects. So assume $\underline{G}^*ab$ and, for reductio, $\neg\exists z\underline{G}az$. Note that the following is an instance of theorem (789.2) with respect to our relation on discernibles $\underline{G}$:

$$(\underline{G}^*ab \,\&\, \forall z(\underline{G}az \rightarrow Fz) \,\&\, Hereditary(F,\underline{G})) \rightarrow Fb$$

By GEN, this theorem holds for all F and so instantiate F to the impossible property $\overline{L}$, where L has been defined on previous occasions as $[\lambda x\, E!x \rightarrow E!x]$, which we know exists:

$$(\underline{G}^*ab \,\&\, \forall z(\underline{G}az \rightarrow \overline{L}z) \,\&\, Hereditary(\overline{L},\underline{G})) \rightarrow \overline{L}b$$

By applying the definition of *Hereditary* and now familiar reasoning, we therefore know:

(ϑ) $(\underline{G}^*ab \,\&\, \forall z(\underline{G}az \rightarrow \overline{L}z) \,\&\, \forall x\forall x'(\underline{G}xx' \rightarrow (\overline{L}x \rightarrow \overline{L}x')) \rightarrow \overline{L}b$

Clearly, the consequent of (ϑ) is false. By definition, nothing exemplifies $\overline{L}$. Hence, one of the following conjuncts of the antecedent of (ϑ) must be false:

(A) $\underline{G}^*ab$

(B) $\forall z(\underline{G}az \rightarrow \overline{L}z)$

(C) $\forall x\forall x'(\underline{G}xx' \rightarrow (\overline{L}x \rightarrow \overline{L}x'))$

But (B) is true: our reductio hypothesis is $\forall z\neg\underline{G}az$ and so *a fortiori*, $\forall z(\underline{G}az \rightarrow \overline{L}z)$. Moreover, (C) is true: since both $\overline{L}x$ and $\overline{L}x'$ are false, $\overline{L}x \rightarrow \overline{L}x'$ is true, and so by the truth of the consequent, $\underline{G}xx' \rightarrow (\overline{L}x \rightarrow \overline{L}x')$. Since this reasoning holds for arbitrary x and x', we've established (C). Hence (A) is false, i.e., $\neg\underline{G}^*ab$, which contradicts our initial assumption. ⋈

(791.2) (Excercise)

(794) (Exercise)

(795.1) Assume $\underline{G}xy$. Then by (789.1), which holds for any strong ancestral, including that of relations on discernibles, $\underline{G}^*xy$. But then by $\vee$I, $\underline{G}^*xy \vee x =_D y$. Hence, $\underline{G}^+xy$, by (794). ⋈

(795.2) Assume Fx, $\underline{G}^{+}xy$, and $Hereditary(F,\underline{G})$. The second assumption implies, by theorem (794), that either $\underline{G}^{*}xy$ or $x =_D y$. If the former, then Fy, by the first and third assumptions and (789.3), which holds for any strong ancestrals, including those of relations on discernibles. If the latter, then by theorem (273.19), $x = y$. So it follows from the first assumption, Fx, that Fy. ⋈

(795.3) Assume $\underline{G}^{+}xy \,\&\, \underline{G}yz$. Then, by theorem (794), $(\underline{G}^{*}xy \vee x =_D y) \,\&\, \underline{G}yz$. So by a variant of (88.6.a), we know either $\underline{G}^{*}xy \,\&\, \underline{G}yz$ or $x =_D y \,\&\, \underline{G}yz$. We show $\underline{G}^{*}xz$ holds in both cases:

> Case 1: $\underline{G}^{*}xy$ and $\underline{G}yz$. But the latter implies $\underline{G}^{*}yz$, by (789.1). Hence by (789.6), $\underline{G}^{*}xz$.
>
> Case 2: $x =_D y$ and $\underline{G}yz$. The former implies $x = y$, by (273.3). Then from $\underline{G}yz$ we know $\underline{G}xz$. So by (789.1), which holds for any strong ancestral, including that of rigid one-to-one relations, it follows that $\underline{G}^{*}xz$. ⋈

(795.4) Assume $\underline{G}^{*}xy$ and $\underline{G}yz$. Then from the first assumption, it follows by $\vee$I that $\underline{G}^{*}xy \vee x =_D y$. So by theorem (794), $\underline{G}^{+}xy$. From this and our second assumption, it follows by (795.3) that $\underline{G}^{*}xz$. So, again by $\vee$I, $\underline{G}^{*}xz \vee x =_D z$, and hence $\underline{G}^{+}xz$, by definition (794). ⋈

(795.5) Assume $\underline{G}xy$ and $\underline{G}^{+}yz$. By the latter and theorem (794), either $\underline{G}^{*}yz \vee y =_D z$. We reason by cases from the two disjuncts. From the first disjunct and our first assumption, it follows that $\underline{G}^{*}xz$, by (789.4), which holds for any strong ancestral. From the second disjunct it follows that $y = z$, by (273.19). So from the first assumption, it follows that $\underline{G}xz$, in which case, $\underline{G}^{*}xz$, by (789.1). ⋈

(795.6) Assume $\underline{G}^{+}xy$ and $\underline{G}^{+}yz$. We reason to the conclusion $\underline{G}^{+}xz$ by cases: $x =_D y$ and $x \neq_D y$.

Case 1. $x =_D y$. Then $x = y$, by (273.19). So from our initial assumption that $\underline{G}^{+}yz$, it follows that $\underline{G}^{+}xz$.

Case 2. $x \neq_D y$. We reason to the conclusion $\underline{G}^{+}xz$ by cases: $y =_D z$ and $y \neq_D z$:

> *Case A*. $y =_D z$. Then $y = z$, by now familiar reasoning. So from our assumption that $\underline{G}^{+}xy$, it follows that $\underline{G}^{+}xz$.
>
> *Case B*. $y \neq_D z$. Then since $\underline{G}^{+}yz$ by assumption, it follows by the basic fact about $\underline{G}^{+}$ (794) that $\underline{G}^{*}yz$. Analogously, since we also know both $x \neq_D y$ (Case 2) and $\underline{G}^{+}xy$ (by assumption), it follows that $\underline{G}^{*}xy$. Now the strong ancestral G^{*} is transitive for any G (789.6). So $\underline{G}^{*}$ is transitive. Hence, $\underline{G}^{*}xz$. So by the main fact about the weak ancestral for relations on discernibles (794), $\underline{G}^{+}xz$. ⋈

(795.7) To avoid clash of variables, we prove $\underline{G}^*ab \rightarrow \exists x(\underline{G}^+ax \,\&\, \underline{G}xb)$, where a and b are any arbitrarily chosen objects. So assume $\underline{G}^*ab$. Note that the following is an instance of (789.2) with respect to our rigid one-to-one relation:

$$(\underline{G}^*ab \,\&\, \forall z(\underline{G}az \rightarrow Fz) \,\&\, Hereditary(F, \underline{G})) \rightarrow Fb$$

Instantiate F in the above to the property $[\lambda y\, \exists x(\underline{G}^+ax \,\&\, \underline{G}xy)]$, which we know exists:

$$(\underline{G}^*ab \,\&\, \forall z(\underline{G}az \rightarrow [\lambda y\, \exists x(\underline{G}^+ax \,\&\, \underline{G}xy)]z) \,\&\, \\ Hereditary([\lambda y\, \exists x(\underline{G}^+ax \,\&\, \underline{G}xy)], \underline{G})) \rightarrow [\lambda y\, \exists x(\underline{G}^+ax \,\&\, \underline{G}xy)]b$$

Since $[\lambda y\, \exists x(\underline{G}^+ax \,\&\, \underline{G}xy)]\downarrow$, applications of β-Conversion and the Rule of Substitution (160.2), reduce the above to:

$$(\underline{G}^*ab \,\&\, \forall z(\underline{G}az \rightarrow \exists x(\underline{G}^+ax \,\&\, \underline{G}xz)) \,\&\, \\ Hereditary([\lambda y\, \exists x(\underline{G}^+ax \,\&\, \underline{G}xy)], \underline{G})) \rightarrow \exists x(\underline{G}^+ax \,\&\, \underline{G}xb)$$

By applying the definition of hereditary (786.1), β-Conversion, and a Rule of Substitution to the result, this becomes:

$$[\underline{G}^*ab \,\&\, \forall z(\underline{G}az \rightarrow \exists x(\underline{G}^+ax \,\&\, \underline{G}xz)) \,\&\, \\ \forall y \forall z(\underline{G}yz \rightarrow (\exists x(\underline{G}^+ax \,\&\, \underline{G}xy) \rightarrow \exists x(\underline{G}^+ax \,\&\, \underline{G}xz)))] \rightarrow \exists x(\underline{G}^+ax \,\&\, \underline{G}xb)$$

Since the consequent is the desired conclusion, it remains only to show the three conjuncts of the antecedent. The first is true by assumption. For the second, assume $\underline{G}az$. Then since $\underline{G}$ is a relation on discernibles, it follows by definition (791.1) that $D!a$. So by (273.18), $a =_D a$. Hence by fact (794), $\underline{G}^+aa$. So, from $\underline{G}^+aa \,\&\, \underline{G}az$, it follows that $\exists x(\underline{G}^+ax \,\&\, \underline{G}xz)$. For the third conjunct, assume $\underline{G}yz$ and $\exists x(\underline{G}^+ax \,\&\, \underline{G}xy)$, by GEN. Suppose c is a witness to the second assumption, so that we know $\underline{G}^+ac \,\&\, \underline{G}cy$. Then by (795.3), it follows that $\underline{G}^*ay$. Thus, by $\vee$I, $\underline{G}^*ay \vee a =_D y$. So $\underline{G}^+ay$, by the main fact about $\underline{G}^+$ (794). Hence we know $\underline{G}^+ay \,\&\, \underline{G}yz$. So by $\exists$I, $\exists x(\underline{G}^+ax \,\&\, \underline{G}xz)$. ⋈

(796.2) Assume *1-1*($\underline{G}$), $\underline{G}xy$, and $\underline{G}^*zy$. By the latter and (795.7), it follows that there is some object, say a, such that $\underline{G}^+za$ and $\underline{G}ay$. Since $\underline{G}$ is a one-to-one relation by hypothesis, it follows that $x = a$ (796.1). So $\underline{G}^+zx$. ⋈

(796.3) Assume *1-1*($\underline{G}$), $\underline{G}xy$ and $\neg\underline{G}^*xx$. For reductio, assume $\underline{G}^*yy$. Now, independently, we know that the following is an instance of (796.2) by setting z in that theorem to y:

$$(\underline{G}xy \,\&\, \underline{G}^*yy) \rightarrow \underline{G}^+yx$$

So, $\underline{G}^+yx$. Similarly, we know that the following is an instance of (795.5) if we set z in that theorem to x:

$$(\underline{G}xy \,\&\, \underline{G}^+yx) \rightarrow \underline{G}^*xx$$

Hence $\underline{G}^* xx$. Contradiction. ⋈

(796.4) Assume $\mathit{1\text{-}1}(\underline{G})$, $\neg\underline{G}^* xx$, and $\underline{G}^+ xy$. Now independently, since $[\lambda z \neg\underline{G}^* zz]\downarrow$, we can instantiate F in (795.2) to $[\lambda z\, \neg\underline{G}^* zz]$ and apply β-Conversion and the Rule of Substitution (160.2) to obtain:

$$(\neg\underline{G}^* xx \,\&\, \underline{G}^+ xy \,\&\, \mathit{Hereditary}([\lambda z\, \neg\underline{G}^* zz], \underline{G})) \rightarrow \neg\underline{G}^* yy$$

Since the consequent is what we want to show, we establish the antecedent. The first two conjuncts of the antecedent are true by assumption. So by definition (786.1), β-Conversion and the Rule of Substitution (160.2), it remains to show:

$$\forall x' \forall y' (\underline{G} x' y' \rightarrow (\neg\underline{G}^* x' x' \rightarrow \neg\underline{G}^* y' y'))$$

So by GEN, it suffices to show:

$$(\vartheta) \quad \underline{G} x' y' \rightarrow (\neg\underline{G}^* x' x' \rightarrow \neg\underline{G}^* y' y')$$

But instantiating $\underline{G}$, x', and y' into the universal closure of (796.3) yields:

$$\mathit{1\text{-}1}(\underline{G}) \rightarrow ((\underline{G} x' y' \,\&\, \neg\underline{G}^* x' x') \rightarrow \neg\underline{G}^* y' y')$$

Since the antecedent holds by assumption, it follows that:

$$(\underline{G} x' y' \,\&\, \neg\underline{G}^* x' x') \rightarrow \neg\underline{G}^* y' y'$$

But by exportation, this is equivalent to (ϑ). ⋈

(797) Consider any relation on discernibles $\underline{G}$ and assume the antecedent of what we have to prove:

$$(\vartheta) \quad Fz \,\&\, \forall x \forall y ((\underline{G}^+ zx \,\&\, \underline{G}^+ zy) \rightarrow (\underline{G} xy \rightarrow (Fx \rightarrow Fy)))$$

The consequent of what we have to prove is $\forall x(\underline{G}^+ zx \rightarrow Fx)$. So by GEN, assume $\underline{G}^+ zx$, to show Fx. We do this by appeal to lemma (795.2). Instantiate the variable F in this lemma to $[\lambda y\, Fy \,\&\, \underline{G}^+ zy]$ (the significance of which we leave as an exercise) and instantiate the variables x and y in the lemma to z and x, respectively. So, by now familiar reasoning, this yields:

$$(\xi) \quad [Fz \,\&\, \underline{G}^+ zz \,\&\, \underline{G}^+ zx \,\&\, \mathit{Hereditary}([\lambda y\, Fy \,\&\, \underline{G}^+ zy], \underline{G})] \rightarrow (Fx \,\&\, \underline{G}^+ zx)$$

So if we can establish the antecedent of (ξ), our desired conclusion, Fx, follows *a fortiori* from the consequent of (ξ). We know the first conjunct of the antecedent of (ξ) is true by the first conjunct of our assumption (ϑ). We know that the second conjunct of the antecedent of (ξ) is true, by the reflexivity of $\underline{G}^+$, which immediately follows from theorem (794) and the reflexivity of $=_D$ on discernible objects (273.30). We know that the third conjunct of the antecedent is true, by further assumption. So it remains to show:

$Hereditary([\lambda y\, Fy \,\&\, \underline{G}^{+}zy], \underline{G}),$

By the definition of *Hereditary* and familiar reasoning, we therefore have to show:

$$\forall x \forall x'(\underline{G}xx' \to ((Fx \,\&\, \underline{G}^{+}zx) \to (Fx' \,\&\, \underline{G}^{+}zx')))$$

Proof. Let a, b be arbitrary objects. Assume $\underline{G}ab$, Fa, and $\underline{G}^{+}za$, to show $Fb \,\&\, \underline{G}^{+}zb$. The second conjunct $\underline{G}^{+}zb$ follows easily: from the facts that $\underline{G}^{+}za$ and $\underline{G}ab$, it follows from (795.3) that $\underline{G}^{*}zb$, which implies $\underline{G}^{+}zb$, by the theorem (794). So it remains to show Fb. Since we now have $\underline{G}^{+}za$, $\underline{G}^{+}zb$, $\underline{G}ab$, and Fa, it follows from the second conjunct of (ϑ) that Fb. ⋈

(801.2) Axiom (800) asserts:

$$[\lambda xy\, \exists F \exists u(Fu \,\&\, Numbers(y, F) \,\&\, Numbers(x, F^{-u}))]{\downarrow}$$

So by the Rule of Identity by Definition (120), it follows from definition (801.1) that:

$$\mathbb{P} = [\lambda xy\, \exists F \exists u(Fu \,\&\, Numbers(y, F) \,\&\, Numbers(x, F^{-u}))]$$

So by (107.1), $\mathbb{P}{\downarrow}$. ⋈

(801.3) (Exercise)

(802.1) Assume $\mathbb{P}xy$. Then by theorem (801.3), this is equivalent, by modally strict reasoning, to:

(A) $\exists F \exists u(Fu \,\&\, Numbers(y, F) \,\&\, Numbers(x, F^{-u}))$

If we show that (A) is necessary, then by (801.3) and a Rule of Substitution, it follows that $\Box Pxy$. Note that to show that (A) is necessary, it suffices, by two applications of the Buridan formula (168.1) and a Rule of Substitution, to show:

(B) $\exists F \exists u \Box(Fu \,\&\, Numbers(y, F) \,\&\, Numbers(x, F^{-u}))$

To find our witnesses, reconsider (A) and let Q and a be such a property and discernible object, respectively, so that we know:

$$Qa \,\&\, Numbers(y, Q) \,\&\, Numbers(x, Q^{-a})$$

Since theorem (573.3) yields $\exists G(Rigidifies(G, Q))$, then let S be such a property, so that we know $Rigidifies(S, Q)$. Hence, we know, by (571.2) that:

(ξ) $Rigid(S)$

(ζ) $\forall x(Sx \equiv Qx)$

If we can show:

(a) $\Box Sa$

(b) $\Box Numbers(y, S)$

(c) $\Box Numbers(x, S^{-a})$

then since a conjunction of necessary truths implies a necessary conjunction of the unnecessitated truths, it follows that S and a are the witnesses we need to establish (B).

To show (a): since we know Qa, (ζ) implies Sa and so by (ξ), $\Box Sa$.

To show (b), it suffices by (769.2) and (ξ), to show $Numbers(y, S)$. But this, in turn, follows from (766) using (ζ) and the fact that *Numbers(y,Q)*.

To show (c), it suffices to show that $Rigidifies(S^{-a}, Q^{-a})$. Specifically, we need only show that S^{-a} is rigid and materially equivalent to Q^{-a}, i.e., that:

(ξ') $\Box\forall z(S^{-a}z \to \Box S^{-a}z)$

(ζ') $\forall z(S^{-a}z \equiv Q^{-a}z)$

(Exercise) ⋈

(802.2) (Exercise)

(802.3) Clearly, $\mathbb{P}\downarrow$. So by (796.1) and GEN, we have to show:

$$\mathbb{P}xz \,\&\, \mathbb{P}yz \to x = y$$

So assume both $\mathbb{P}xz$ and $\mathbb{P}yz$. Then by theorem (801.3), these assumptions imply, respectively, that there are properties and discernible objects, say R, Q, a, b, such that:

(ϑ) $Qa \,\&\, Numbers(z, Q) \,\&\, Numbers(x, Q^{-a})$

(ξ) $Rb \,\&\, Numbers(z, R) \,\&\, Numbers(y, R^{-b})$

The second conjuncts of (ϑ) and (ξ) jointly yield $Q \approx_D R$, by (764.2). Since we also know Qa and Rb, it follows by lemma (753) that $Q^{-a} \approx_D R^{-b}$. But, separately, the third conjuncts of (ϑ) and (ξ) jointly imply $x = y \equiv Q^{-a} \approx_D R^{-b}$, by a Principle Underlying Hume's Principle (765.1). Hence $x = y$. ⋈

(802.4) Assume both $\mathbb{P}xy$ and $\mathbb{P}xz$. Then given the necessary and sufficient conditions for *predecessor* (801.3), these assumptions imply, respectively, that there are properties and discernible objects, say Q, R, a, b, such that:

(ϑ) $Qa \,\&\, Numbers(y, Q) \,\&\, Numbers(x, Q^{-a})$

(ξ) $Rb \,\&\, Numbers(z, R) \,\&\, Numbers(x, R^{-b})$

Now the third conjuncts of (ϑ) and (ξ) jointly imply $Q^{-a} \approx_D R^{-b}$, by (764.2). Since we also know Qa and Rb, it follows by lemma (754) that $Q \approx_D R$. But independently, the second conjuncts of (ϑ) and (ξ) jointly imply $y = z \equiv Q \approx_D R$, by a Principle Underlying Hume's Principle (765.1). Hence $y = z$. ⋈

(803.1) By (273.5), $\exists x D!x$. Let a be such an object, so that we know $D!a$. Then by (273.34), a's haecceity exists, i.e., $[\lambda x\, x = a]\downarrow$. So by (771.2), $\#[\lambda x\, x = a]\downarrow$. Then we can establish our theorem by showing 0 and $\#[\lambda x\, x = a]$ are witnesses, i.e., that $\mathbb{P}0\#[\lambda x\, x = a]$. By (801.3), we have to show the following:

(ϑ) $\exists F \exists u(Fu \;\&\; Numbers(\#[\lambda x\, x = a], F) \;\&\; Numbers(0, F^{-u}))$

Now by theorem (774.2), $Numbers(\#[\lambda x\, x = a], [\lambda z\, \mathcal{A}[\lambda x\, x = a]z])$. So to show that the witnesses to (ϑ) are, respectively, $[\lambda z\, \mathcal{A}[\lambda x\, x = a]z]$ and a, it remains to show:

(i) $[\lambda z\, \mathcal{A}[\lambda x\, x = a]z]a$

(ii) $Numbers(0, [\lambda z\, \mathcal{A}[\lambda x\, x = a]z]^{-a})$

(i) Since $a = a$ (117.1), it follows that $\mathcal{A}a = a$ (175.1). But since $[\lambda x\, x = a]\downarrow$, strengthened β-Conversion implies that $[\lambda x\, x = a]a \equiv a = a$, as a modally strict theorem. Hence, by a Rule of Substitution, $\mathcal{A}[\lambda x\, x = a]a$. And from this last fact and the facts that $[\lambda z\, \mathcal{A}[\lambda x\, x = a]z]\downarrow$ and $a\downarrow$, it follows by Rule $\overleftarrow{\beta}$C (184.2.a) that $[\lambda z\, \mathcal{A}[\lambda x\, x = a]z]a$.

(ii) To show $Numbers(0, [\lambda z\, \mathcal{A}[\lambda x\, x = a]z]^{-a})$, it suffices to show, by (784.1):

$$\neg\exists u([\lambda z\, \mathcal{A}[\lambda x\, x = a]z]^{-a}u)$$

For reductio, suppose that there is such a discernible object, say b, so that we know $[\lambda z\, \mathcal{A}[\lambda x\, x = a]z]^{-a}b$. Then by definition (752.2), we know:

$$[\lambda y\, [\lambda z\, \mathcal{A}[\lambda x\, x = a]z]y \;\&\; y \neq a]b$$

By (752.1), $[\lambda y\, [\lambda z\, \mathcal{A}[\lambda x\, x = a]z]y \;\&\; y \neq a]$ exists, since a is discernible. So by β-Conversion:

(ξ) $[\lambda z\, \mathcal{A}[\lambda x\, x = a]z]b \;\&\; b \neq a$

The first conjunct of (ξ) implies $\mathcal{A}[\lambda x\, x = a]b$, also by β-Conversion. Since, as a modally strict theorem, we know $[\lambda x\, x = a]b \equiv b = a$, it follows by a Rule of Substitution that $\mathcal{A}b = a$. Hence $b = a$ (175.1), which contradicts the second conjunct of (ξ). ⋈

(803.2) Assume $NaturalCardinal(x)$ and $x \neq 0$. We want to show $\exists y \mathbb{P}yx$, i.e.,

(ϑ) $\exists y \exists F \exists u(Fu \;\&\; Numbers(x, F) \;\&\; Numbers(y, F^{-u}))$

From the first assumption, we know, by definition (777), that $\exists G(x = \#G)$, and so by (778.2), it follows that $\exists G(\mathit{Numbers}(x, G))$. Suppose P is such a property, so that we know $\mathit{Numbers}(x, P)$. This and the fact that $x \neq 0$ imply, by (784.2), that $\exists u Pu$. Suppose a is such a discernible object, so that we know Pa. Then it follows, by (752.2), that $[\lambda z\, Pz \,\&\, z \neq a]\downarrow$. Hence $P^{-a}\downarrow$. So $\exists y \mathit{Numbers}(y, P^{-a})$, by (763.1). Let b be such an object, so that we know $\mathit{Numbers}(b, P^{-a})$. Thus, assembling what we know:

$$Pa \,\&\, \mathit{Number}(x, P) \,\&\, \mathit{Numbers}(b, P^{-a})$$

Then (ϑ) follows from this last result by 3 applications of $\exists$I. ⋈

(803.3) Assume $\mathit{NaturalCardinal}(x)$. Since Zero is a natural cardinal (783), we show $D!x$ by disjunctive syllogism from $x = 0 \vee x \neq 0$.

(a) $x = 0$. Note that as part of the reasoning in (803.1), we established that $\exists y \mathbb{P}0y$. Let b be such an object, so that we know $\mathbb{P}0b$. Then since $[\lambda z \mathbb{P}zb]\downarrow$ and $0\downarrow$, it follows that $[\lambda z\, \mathbb{P}zb]0$. To show $D!0$, we have to show, by GEN, $y \neq 0 \rightarrow \exists F \neg(Fy \equiv F0)$ (273.3). So assume $y \neq 0$. Suppose, for reductio, $\neg\exists F \neg(Fy \equiv F0)$, i.e., $\forall F(Fy \equiv F0)$. Then $[\lambda z\, \mathbb{P}zb]y$, and so $\mathbb{P}yb$. But $\mathbb{P}$ is a 1-1 relation (802.3), and so by the definition of 1-1 relations (796.1) and the fact that $\mathbb{P}\downarrow$ (801.2), we may infer $0 = y$ from the previously established facts that $\mathbb{P}0b$ and $\mathbb{P}yb$. But then, since $x = 0$, it follows that $x = y$, which contradicts our hypothesis that $y \neq x$.

(b) $x \neq 0$. Then since x is a natural cardinal, it follows by (803.2), that $\exists y \mathbb{P}yx$. For let c be such an object, so that we know $\mathbb{P}cx$. Then since $[\lambda z\, \mathbb{P}cz]\downarrow$ and $x\downarrow$, it follows that $[\lambda z\, \mathbb{P}cz]x$. To show $D!x$, we have to show, by GEN, $y \neq x \rightarrow \exists F \neg(Fy \equiv Fx)$ (273.3). So assume $y \neq x$. Suppose, for reductio, $\neg\exists F \neg(Fy \equiv Fx)$, i.e., $\forall F(Fy \equiv Fx)$. Then $[\lambda z \mathbb{P}cz]y$, and hence $\mathbb{P}cy$. But $\mathbb{P}$ is a functional relation (802.4) and so we may infer from $\mathbb{P}cx$ and $\mathbb{P}cy$ that $x = y$, which contradicts our assumption that $y \neq x$. ⋈

(803.4) Assume $\mathbb{P}xy$. Then by theorem (801.3), it follows that there is a property, say Q, and a discernible object, say b, such that:

$$Qb \,\&\, \mathit{Numbers}(x, Q) \,\&\, \mathit{Numbers}(y, Q^{-b})$$

The second and third conjuncts imply, respectively, by (778.1), that x and y are natural cardinals. ⋈

(803.5) This follows directly from (803.3) and (803.4). ⋈

(803.6) By (273.13), we know $[\lambda x\, D!x \& \mathit{Numbers}(x, F)]\downarrow$. By (778.1) and (803.3), it is straightforward to establish:

$$\Box\forall x(D!x \,\&\, \mathit{Numbers}(x, F) \equiv \mathit{Numbers}(x, F))$$

Hence, by an instance of axiom (49), we may conclude $[\lambda x\, Numbers(x,F)]{\downarrow}$. ⋈

(804.1) By theorem (786.2) we know:

$$[\lambda xy\ \forall F((\forall z(Gxz \rightarrow Fz)\ \&\ Hereditary(F,G)) \rightarrow Fy)]{\downarrow}$$

Hence, by the Rule of Identity by Definition (120), it follows from definition (787) that:

$$G^{*} = [\lambda xy\ \forall F((\forall z(Gxz \rightarrow Fz)\ \&\ Hereditary(F,G)) \rightarrow Fy)]$$

Since this is a theorem, it follows by GEN that:

$$\forall G(G^{*} = [\lambda xy\ \forall F((\forall z(Gxz \rightarrow Fz)\ \&\ Hereditary(F,G)) \rightarrow Fy)])$$

Since $\mathbb{P}{\downarrow}$ (801.2), it follows that:

$$\mathbb{P}^{*} = [\lambda xy\ \forall F((\forall z(\mathbb{P}xz \rightarrow Fz)\ \&\ Hereditary(F,\mathbb{P})) \rightarrow Fy)] \qquad ⋈$$

(804.2) – **(804.3)** (Exercises)

(804.4) By (571).1, we have to show:

$$\mathbb{P}^{*}{\downarrow}\ \&\ \Box\forall x\forall y(\mathbb{P}^{*}xy \rightarrow \Box\mathbb{P}^{*}xy)$$

The first conjunct is just theorem (804.2), and to show the second conjunct, it suffices by GEN and RN to show $\mathbb{P}^{*}xy \rightarrow \Box\mathbb{P}^{*}xy$. Our proof strategy begins by considering a variant of $\mathbb{P}^{*}$, designated as $\mathbb{P}^{\star}$, which relates x and y whenever x is a strong ancestor of y w.r.t. *rigid* properties that are hereditary on $\mathbb{P}$, i.e.,

$$\mathbb{P}^{\star}xy \equiv_{df} \\ \forall F\big(Rigid(F) \rightarrow [(\forall z(Pxz \rightarrow Fz)\ \&\ \forall x'\forall y'(Px'y' \rightarrow (Fx' \rightarrow Fy'))) \rightarrow Fy]\big)$$

With this notion, our strategy then is to proceed as follows:

(ϑ) Show: $\mathbb{P}^{\star}xy \rightarrow \Box\mathbb{P}^{\star}xy$

(ξ) Show: $\Box(\mathbb{P}^{*}xy \equiv \mathbb{P}^{\star}xy)$

Our theorem is then implied by these two claims, by the following reasoning. By (158.6), we know that $\Box(\varphi \equiv \psi) \rightarrow (\Box\varphi \equiv \Box\psi)$. So ($\xi$) implies:

(ζ) $\Box\mathbb{P}^{*}xy \equiv \Box\mathbb{P}^{\star}xy$

Hence, we may argue as follows:

$\mathbb{P}^{*}xy$	$\rightarrow$	$\mathbb{P}^{\star}xy$	by (ξ) and the T schema
	$\rightarrow$	$\Box\mathbb{P}^{\star}xy$	by (ϑ)
	$\rightarrow$	$\Box\mathbb{P}^{*}xy$	by (ζ)

This establishes $\mathbb{P}^{*}xy \to \Box\mathbb{P}^{*}xy$ by modally strict reasoning from no assumptions. So it remains only to show (ϑ) and (ξ).

Proof of (ϑ). Assume $\mathbb{P}^{\star}xy$ and, for reductio, $\neg\Box\mathbb{P}^{\star}xy$, i.e., that $\Diamond\neg\mathbb{P}^{\star}xy$. Then, from these assumptions, we know, respectively, by definition of $\mathbb{P}^{\star}xy$, that:

(A) $\forall F\big(Rigid(F) \to [(\forall z(\mathbb{P}xz \to Fz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Fx' \to Fy'))) \to Fy]\big)$

(B) $\Diamond\neg\forall F\big(Rigid(F) \to [(\forall z(\mathbb{P}xz \to Fz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Fx' \to Fy'))) \to Fy]\big)$

From (B), it follows that:

$$\Diamond\exists F\big(Rigid(F) \,\&\, \neg[(\forall z(\mathbb{P}xz \to Fz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Fx' \to Fy'))) \to Fy]\big)$$

So by BF$\Diamond$:

$$\exists F\Diamond\big(Rigid(F) \,\&\, \neg[(\forall z(\mathbb{P}xz \to Fz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Fx' \to Fy'))) \to Fy]\big)$$

Let Q be such a property so that we know:

(C) $\Diamond\big(Rigid(Q) \,\&\, \neg[(\forall z(\mathbb{P}xz \to Qz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Qx' \to Qy'))) \to Qy]\big)$

Now (C) has the form $\Diamond$(D), where:

(D) $Rigid(Q) \,\&\, \neg[(\forall z(\mathbb{P}xz \to Qz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Qx' \to Qy'))) \to Qy]$

Now independently we can establish that (D) $\to \Box$(D) by modally strict means:

Proof. Assume (D). Then we know:

(i) $Rigid(Q)$, i.e., $\Box\forall x(Qx \to \Box Qx)$

(ii) $\neg[(\forall z(\mathbb{P}xz \to Qz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Qx' \to Qy'))) \to Qy]$

Now if we can show that both (i) and (ii) are necessary, then by (158.3), their conjunction is necessary, i.e., we may infer $\Box$(D). But (i) is necessary, by the 4 schema. To see that (ii) is necessary, note first that by the modally strict equivalence $\neg[(\varphi \,\&\, \psi) \to \chi] \equiv (\varphi \,\&\, \psi \,\&\, \neg\chi)$, (ii) implies:

(E) $\forall z(\mathbb{P}xz \to Qz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Qx' \to Qy')) \,\&\, \neg Qy$

But since $\mathbb{P}$ is a rigid relation (802.2) and Q is, by (i), a rigid property, all three conjuncts of (E) are modally collapsed – for each conjunct φ in (E), we know, by the theorems in (171), that $\Box(\varphi \to \Box\varphi)$. For example, the rigidity of $\mathbb{P}$ and Q, together with theorems (171.2) and (171.6), imply that each of the first two conjuncts of (E) imply their own necessity (exercise). And the rigidity of Q and theorem (171.1) imply that the third conjunct of (E) implies its own necessity (exercise). So we can derive that each of the conjuncts of (E) is necessary, to obtain:

$$\Box\forall z(\mathbb{P}xz \to Qz) \,\&\, \Box\forall x'\forall y'(\mathbb{P}x'y' \to (Qx' \to Qy')) \,\&\, \Box\neg Qy$$

From this, by (158.3), it follows that:

$$\Box(\forall z(\mathbb{P}xz \to Qz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Qx' \to Qy')) \,\&\, \neg Qy)$$

So by the modally strict equivalence mentioned above and a Rule of Substitution:

$$\Box\neg[(\forall z(\mathbb{P}xz \to Qz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Qx' \to Qy'))) \to Qy]$$

i.e., (ii) is necessary. Hence, $\Box$(D). ⋈

Since we've established (D) $\to \Box$(D) by modally strict means from no assumptions, it follows by RN that $\Box$((D) $\to \Box$(D)). So from this and (C), which is $\Diamond$(D), it follows by (158.13) that $\Diamond\Box$(D), i.e.,

$$\Diamond\Box\big(Rigid(Q) \,\&\, \neg[(\forall z(\mathbb{P}xz \to Qz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Qx' \to Qy'))) \to Qy]\big)$$

But by the B$\Diamond$ schema (165.4), this collapses to (D), which contradicts (A).

Proof of (ξ). By Rule RN, we have to prove $\mathbb{P}^*xy \equiv \mathbb{P}^\star xy$ by modally strict means from no assumptions. We prove both directions separately. ($\to$) Assume $\mathbb{P}^*xy$. Assume, for reductio, $\neg\mathbb{P}^\star xy$. These assumptions imply, respectively:

(F) $\forall F\big([\forall z(\mathbb{P}xz \to Fz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Fx' \to Fy'))] \to Fy\big)$

(G) $\neg\forall F\big(Rigid(F) \to [(\forall z(\mathbb{P}xz \to Fz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Fx' \to Fy'))) \to Fy]\big)$

Now (G) implies:

$$\exists F\big(Rigid(F) \,\&\, \neg[(\forall z(\mathbb{P}xz \to Fz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Fx' \to Fy'))) \to Fy]\big)$$

Let Q be such a property, so that we know:

$$Rigid(Q) \,\&\, \neg[(\forall z(\mathbb{P}xz \to Qz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Qx' \to Qy'))) \to Qy]$$

But the second conjunct contradicts the result of instantiating (F) to Q.

($\leftarrow$) Assume $\mathbb{P}^\star xy$, and for reductio, $\neg\mathbb{P}^*xy$. These assumptions imply, respectively:

(H) $\forall F\big(Rigid(F) \to [(\forall z(\mathbb{P}xz \to Fz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Fx' \to Fy'))) \to Fy]\big)$

(J) $\neg\forall F\big([\forall z(\mathbb{P}xz \to Fz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Fx' \to Fy'))] \to Fy\big)$

Since (J) implies:

$$\exists F\neg\big([\forall z(\mathbb{P}xz \to Fz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Fx' \to Fy'))] \to Fy\big)$$

let R be such a property, so that we know:

(K) $\neg([\forall z(\mathbb{P}xz \to Rz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Rx' \to Ry'))] \to Ry)$

This implies:

(L) $\forall z(\mathbb{P}xz \to Rz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Rx' \to Ry')) \,\&\, \neg Ry$

Now by (573.3), $\exists F(\mathit{Rigidifies}(F,R))$. Suppose S is such a property, so that we know $\mathit{Rigidifies}(S,R)$. Then by definition (571.2):

$$\mathit{Rigid}(S) \,\&\, \forall x(Sx \equiv Rx)$$

Since S is rigid, it follows from (H) that:

(M) $[\forall z(\mathbb{P}xz \to Sz) \,\&\, \forall x'\forall y'(\mathbb{P}x'y' \to (Sx' \to Sy'))] \to Sy$

But we know from (L) that $\forall z(\mathbb{P}xz \to Rz)$. Since R and S are materially equivalent, it follows by a variant of (99.10) that $\forall z(\mathbb{P}xz \to Sz)$. And we also know from (L) that $\forall x'\forall y'(\mathbb{P}x'y' \to (Rx' \to Ry'))$. But again, since R and S are materially equivalent, it follows that $\forall x'\forall y'(\mathbb{P}x'y' \to (Sx' \to Sy'))$ (exercise). We've therefore established both conjuncts of the antecedent of (M). Hence Sy, and since S and R are materially equivalent, Ry, which contradicts the third conjunct of (L). ⋈

(805.1) Suppose, for reductio, that something, say a, is such that $\mathbb{P}a0$. Then, by the theorem governing $\mathbb{P}$ (801.3), it follows that there is a property, say Q, and a discernible object, say b, such that:

$$Qb \,\&\, \mathit{Numbers}(0,Q) \,\&\, \mathit{Numbers}(a,Q^{-b})$$

From Qb it follows that $\exists uQu$. But from $\mathit{Numbers}(0,Q)$, it follows by (784.1) that $\neg\exists uQu$. Contradiction. ⋈

(805.2) Assume, for reductio, $\exists x\mathbb{P}^*x0$. Suppose a is such an object, so that we know $\mathbb{P}^*a0$. By (789.5), it follows that $\exists x\mathbb{P}x0$. But this contradicts (805.1). ⋈

(805.3) (Exercise)

(806.1) – **(806.2)** (Exercises)

(806.3) By (571).1, we have to show:

$$\mathbb{P}^+\!\downarrow \,\&\, \Box\forall x\forall y(\mathbb{P}^+xy \to \Box\mathbb{P}^+xy)$$

The first conjunct is just theorem (806.1), and to show the second conjunct, it suffices by GEN and RN to show $\mathbb{P}^+xy \to \Box\mathbb{P}^+xy$. So assume $\mathbb{P}^+xy$. Then by fact (806.2), $\mathbb{P}^*xy \vee x =_D y$. But by (804.4), $\mathbb{P}^*xy \to \Box\mathbb{P}^*xy$. And by (273.21), $x =_D y \to \Box x =_D y$. Hence $\Box\mathbb{P}^*xy \vee \Box x =_D y$. So by (158.15), $\Box(\mathbb{P}^*xy \vee x =_D y)$. Hence $\Box\mathbb{P}^+xy$, by fact (806.2). ⋈

(807.2) – **(807.3)** (Exercises)

(808) As an instance of (807.3), we know that $\mathbb{N}0 \equiv \mathbb{P}^{+}00$. So it suffices to show $\mathbb{P}^{+}00$. By theorem (806.2), it suffices to show $\mathbb{P}^{*}00 \vee 0 =_D 0$, and *a fortiori*, to show $0 =_D 0$. But $D!0$ by (783) and (803.3). So, by (273.18), $0 =_D 0$. ⋈

(809.1) Assume $\mathbb{N}x$. Then by (807.3), $\mathbb{P}^{+}0x$. But by (806.3), we know that $\mathbb{P}$ is rigid. Hence $\Box\mathbb{P}^{+}0x$. But then by the fact (807.3) is a modally strict theorem and a Rule of Substitution it follows that $\Box\mathbb{N}x$. ⋈

(809.2) (Exercise)

(810) By (805.1), nothing is a predecessor of Zero. *A fortiori*, no natural number is a predecessor of Zero. ⋈

(811) Though the proof given in the text prior to the statement of the theorem is perfectly adequate, those interested in system implementation will note that the following is required. By the one-to-one character of $\mathbb{P}$ (802.3), we know:

$$\mathbb{P}xz \,\&\, \mathbb{P}yz \rightarrow x = y$$

A fortiori:

$$(\mathbb{N}x \,\&\, \mathbb{N}y \,\&\, \mathbb{N}z) \rightarrow (\mathbb{P}xz \,\&\, \mathbb{P}yz \rightarrow x = y)$$

So by GEN,

$$\forall x\forall y\forall z((\mathbb{N}x \,\&\, \mathbb{N}y \,\&\, \mathbb{N}z) \rightarrow (\mathbb{P}xz \,\&\, \mathbb{P}yz \rightarrow x = y))$$

Hence, by the conventions for restricted variables, this becomes:

$$\forall n\forall m\forall k(\mathbb{P}nk \,\&\, \mathbb{P}mk \rightarrow n = m) \qquad ⋈$$

(812) By GEN, it suffices to show $F0 \,\&\, \forall n\forall m(\mathbb{P}nm \rightarrow (Fn \rightarrow Fm)) \rightarrow \forall nFn$. Now since $\mathbb{P}$ is a relation on discernibles, we can instantiate it for $\underline{G}$ in the Principle of Generalized Induction (797). If we do so and simultaneously substitute Zero for z, we obtain the following instance:

$$[F0 \,\&\, \forall x\forall y((\mathbb{P}^{+}0x \,\&\, \mathbb{P}^{+}0y) \rightarrow (\mathbb{P}xy \rightarrow (Fx \rightarrow Fy)))] \rightarrow \forall x(\mathbb{P}^{+}0x \rightarrow Fx)$$

By the main theorem governing $\mathbb{N}$ (807.3) and the Rule of Substitution, this reduces to:

$$[F0 \,\&\, \forall x\forall y((\mathbb{N}x \,\&\, \mathbb{N}y) \rightarrow (\mathbb{P}xy \rightarrow (Fx \rightarrow Fy)))] \rightarrow \forall x(\mathbb{N}x \rightarrow Fx)$$

By employing our restricted variables n and m, this can be written:

$$[F0 \,\&\, \forall n\forall m(\mathbb{P}nm \rightarrow (Fn \rightarrow Fm))] \rightarrow \forall nFn \qquad ⋈$$

(813.1) Assume $\mathbb{N}x$. Then, by (807.3), $\mathbb{P}^{+}0x$. We then have two cases. If $x=0$, then *NaturalCardinal*(x), by (783). If $x \neq 0$, then it follows that $\mathbb{P}^{*}0x$, by (806.2) and the fact that $x \neq 0 \rightarrow x \neq_D 0$ (273.19). By (789.7), it follows *a fortiori* that $\exists z \mathbb{P}zx$. Let a be such an object, so that we know $\mathbb{P}ax$. Then by (803.4) again *a fortiori* that *NaturalCardinal*(x). ⋈

(813.2) (Exercise)

(814.1) Assume the antecedent $\mathbb{P}nx$. Since n is, by hypothesis, a natural number, we know by (807.3) that $\mathbb{P}^{+}0n$. Independently, since $\mathbb{P}$ is a relation on discernibles, we can instantiate it for $\underline{G}$ in (795.3) to produce the following instance:

$$(\mathbb{P}^{+}0n \,\&\, \mathbb{P}nx) \rightarrow \mathbb{P}^{*}0x$$

So $\mathbb{P}^{*}0x$. Hence, by $\vee$I and theorem (806.2), it follows that $\mathbb{P}^{+}0x$. So $\mathbb{N}x$, by (807.3). ⋈

(814.2) For convenience, we prove $\mathbb{P}^{*}ny \rightarrow \mathbb{N}y$. So assume $\mathbb{P}^{*}ny$. Then by the main theorem governing the strong ancestral of *predecessor* (804.3):

$$\forall F((\forall z(\mathbb{P}nz \rightarrow Fz) \,\&\, \forall x \forall x'(\mathbb{P}xx' \rightarrow (Fx \rightarrow Fx'))) \rightarrow Fy)$$

If we instantiate this to $\mathbb{N}$, then we know:

$$(\vartheta) \quad (\forall z(\mathbb{P}nz \rightarrow \mathbb{N}z) \,\&\, \forall x \forall x'(\mathbb{P}xx' \rightarrow (\mathbb{N}x \rightarrow \mathbb{N}x'))) \rightarrow \mathbb{N}y$$

Since we're trying to show $\mathbb{N}y$, we simply have to establish both conjuncts of the antecedent of (ϑ). For the first conjunct, assume $\mathbb{P}nz$. But then $\mathbb{N}z$ follows by (814.1). So it remains to show the second conjunct of (ϑ). By GEN and conditional proof, assume $\mathbb{P}xx'$ and $\mathbb{N}x$, to show $\mathbb{N}x'$. But these two assumptions imply $\mathbb{N}x'$, also by (814.1). ⋈

(814.3) Assume $\mathbb{P}^{+}nx$. Then by theorem (806.2), it follows that $\mathbb{P}^{*}nx \vee n =_D x$. Reason by cases from this disjunction. If $\mathbb{P}^{*}nx$, then by (814.2), $\mathbb{N}x$. If $n =_D x$, then by (273.19), $n=x$. So $\mathbb{N}x$, given that $\mathbb{N}n$ (n is a restricted variable ranging over natural numbers). ⋈

(815) Assume $\mathbb{P}xn$. Since n is by hypothesis a natural number, we know by theorem (807.3) that $\mathbb{P}^{+}0n$. Hence $\mathbb{P}^{*}0n \vee 0 =_D n$, by theorem (806.2). It can't be that $0 =_D n$ for that implies $0=n$ (273.19), which would imply $\mathbb{P}x0$, contradicting (805.1). Then $\mathbb{P}^{*}0n$. Now, independently, it follows from an instance of (796.2) and the fact that $\mathbb{P}$ is both a relation on discernibles and one-to-one (802.3), that $(\mathbb{P}xn \,\&\, \mathbb{P}^{*}0n) \rightarrow \mathbb{P}^{+}0x$. Hence, $\mathbb{P}^{+}0x$, i.e., $\mathbb{N}x$. ⋈

(815.2) Assume $\mathbb{P}^{*}xn$.

(815.3) Assume $\mathbb{P}^{+}xn$. Then by definition, $\mathbb{P}^{*}xn \vee x=n$. If the former, then by **(815.2)**, $\mathbb{N}x$. If the latter, then since n is by hypothesis a number, so is x. ⋈

(816) (Exercise)

(817.1) Assume $\mathbb{N}b$, where b is an arbitrarily chosen object. By Rule ∀I (96), it suffices to show: $\neg\mathbb{P}^*bb$. Now theorem (795.2), concerning the weak ancestrals of relations on discernibles is:

$$(Fx \,\&\, \underline{G}^+(x,y) \,\&\, \mathit{Hereditary}(F,\underline{G})) \rightarrow Fy$$

So where x is 0, y is b, F is $[\lambda z\, \neg\mathbb{P}^*zz]$, and $\underline{G}$ is $\mathbb{P}$, we know:

$$([\lambda z\, \neg\mathbb{P}^*zz]0 \,\&\, \mathbb{P}^+(0,b) \;\; \&\, \mathit{Hereditary}([\lambda z\, \neg\mathbb{P}^*zz],\mathbb{P})) \rightarrow [\lambda z\, \neg\mathbb{P}^*zz]b$$

By applications of β-Conversion and a Rule of Substitution, this reduces to:

$$(\neg\mathbb{P}^*00 \,\&\, \mathbb{P}^+0b \;\; \&\, \mathit{Hereditary}([\lambda z\, \neg\mathbb{P}^*zz],\mathbb{P})) \rightarrow \neg\mathbb{P}^*bb$$

So to establish $\neg\mathbb{P}^*bb$, we need to show each of the following conjuncts of the antecedent:

(ϑ) $\neg\mathbb{P}^*00$

(ξ) $\mathbb{P}^+0b$

(ζ) $\mathit{Hereditary}([\lambda z\, \neg\mathbb{P}^*zz],\mathbb{P})$

(ϑ) follows immediately from theorem (805.2), which asserts $\neg\exists x\mathbb{P}^*x0$. ($\xi$) follows from our assumption that $\mathbb{N}b$ and the definition of $\mathbb{N}$. So to prove (ζ), we have to show, by definition (786.1):

$$\mathbb{P}{\downarrow} \,\&\, [\lambda z\, \neg\mathbb{P}^*zz]{\downarrow} \,\&\, \forall x\forall y(\mathbb{P}xy \rightarrow ([\lambda z\, \neg\mathbb{P}^*zz]x \rightarrow [\lambda z\, \neg\mathbb{P}^*zz]y))$$

The first conjunct is theorem (801.2); the second is easily established (exercise). So it remains to show the third, which by applications of β-Conversion and a Rule of Substitution, reduces to $\forall x\forall y(\mathbb{P}xy \rightarrow (\neg\mathbb{P}^*xx \rightarrow \neg\mathbb{P}^*yy))$. So by GEN, we have to show:

$$\mathbb{P}xy \rightarrow (\neg\mathbb{P}^*xx \rightarrow \neg\mathbb{P}^*yy)$$

Assume $\mathbb{P}xy$ and $\neg\mathbb{P}^*xx$. Now by (802.2), (802.3), and (803.5), $\mathbb{P}$ is a 1-1 relation on discernibles. So by instantiating $\mathbb{P}$, x, and y into the universal generaliation of theorem (796.3) and detaching the consequent, we may infer:

$$(\mathbb{P}xy \,\&\, \neg\mathbb{P}^*xx) \rightarrow \neg\mathbb{P}^*yy$$

Hence $\neg\mathbb{P}^*yy$. ⋈

(817.2) We take $\mathbb{N}x$ and $\mathbb{P}yx$ as global assumptions, to show:

$$\mathit{Numbers}(z,[\lambda z\, \mathbb{P}^+zy]) \equiv \mathit{Numbers}(z,[\lambda z\, \mathbb{P}^+zx]^{-x})$$

By (766), it suffices to show $[\lambda z\, \mathbb{P}^{+}zy] \equiv_D [\lambda z\, \mathbb{P}^{+}zx]^{-x}$. So, by definition (756) and GEN, we show:

$$[\lambda z\, \mathbb{P}^{+}zy]u \equiv [\lambda z\, \mathbb{P}^{+}zx]^{-x}u$$

Note that since x is a natural number (by hypothesis) and so a discernible object (813.2), $[\lambda z \mathbb{P}^{+}zx]^{-x}$ is properly defined by an instance of definition (752.2). So, in the line displayed above, we can replace $[\lambda z\, \mathbb{P}^{+}zx]^{-x}$ by it definiens $[\lambda z\, [\lambda z\, \mathbb{P}^{+}zx]z \,\&\, z \neq x]$ and apply β-Conversion to both sides of the resulting biconditional. So it suffices to show:

$$\mathbb{P}^{+}uy \equiv [\lambda z\, \mathbb{P}^{+}zx]u \,\&\, u \neq x$$

And by applying β-Conversion and a Rule of Substitution to the right side, we have to show:

$$\mathbb{P}^{+}uy \equiv \mathbb{P}^{+}ux \,\&\, u \neq x$$

($\rightarrow$) Assume $\mathbb{P}^{+}uy$. From this, our global assumption $\mathbb{P}yx$, and the fact that $\mathbb{P}$ is a relation on discernibles, it follows by an appropriate instance of (795.5) that $\mathbb{P}^{*}ux$. Hence $\mathbb{P}^{+}ux$ (806.2). To see that $u \neq x$, suppose $u = x$, for reductio. Then from the previously established fact that $\mathbb{P}^{*}ux$, it follows that $\mathbb{P}^{*}xx$, which contradicts (817.1) given that $\mathbb{N}x$.

($\leftarrow$) Let $\mathbb{P}^{+}ux$ and $u \neq x$ be our local assumptions. Suppose, for reductio, $\neg\mathbb{P}^{+}uy$. From our second local assumption it follows that $u \neq_D x$, by (273.19) and (273.25), and from this and the first local assumption it follows that $\mathbb{P}^{*}ux$ (806.2). But since $\mathbb{P}$ is a 1-1 relation on discernibles, by (802.2) and (802.3), the following is a consequence of (796.2), simultaneously substituting y for x, x for y and u for z:

$$(\mathbb{P}yx \,\&\, \mathbb{P}^{*}ux) \rightarrow \mathbb{P}^{+}uy$$

This is equivalent to:

$$(\mathbb{P}yx \,\&\, \neg\mathbb{P}^{+}uy) \rightarrow \neg\mathbb{P}^{*}ux)$$

But $\mathbb{P}yx$ is a global assumption and $\neg\mathbb{P}^{+}uy$ is our reductio hypothesis. Hence $\neg\mathbb{P}^{*}ux$. Contradiction. ⋈

(817.3) Assume $\mathbb{N}x$ and $\mathbb{P}(y,x)$. These imply, by (817.2) and (764.2), that:

$$(\vartheta)\quad [\lambda z\, \mathbb{P}^{+}zy] \approx_D [\lambda z\, \mathbb{P}^{+}zx]^{-x}$$

Note that from the fact that $\mathbb{P}^{+}$ is rigid, it follows that:

$$Rigid([\lambda z\, \mathbb{P}^{+}zy])$$

$$Rigid([\lambda z\, \mathbb{P}^{+}zx]^{-x})$$

(We leave the proofs as exercises.) By (775), these imply that a modally strict version of Hume's Principle holds, so that we know:

$$\#[\lambda z\, \mathbb{P}^{+}zy] = \#[\lambda z\, \mathbb{P}^{+}zx]^{-x} \equiv [\lambda z\, \mathbb{P}^{+}zy] \approx_{D} [\lambda z\, \mathbb{P}^{+}zx]^{-x}$$

So by (ϑ):

$$\#[\lambda z\, \mathbb{P}^{+}zy] = \#[\lambda z\, \mathbb{P}^{+}zx]^{-x} \qquad \bowtie$$

(817.4) We want to show:

$$[\lambda x\, \exists y(Numbers(y, [\lambda z\, \mathbb{P}^{+}zx]) \,\&\, \mathbb{P}xy)]\downarrow$$

Note that the matrix of the above λ-expression:

$$\exists y(Numbers(y, [\lambda z\, \mathbb{P}^{+}zx]) \,\&\, \mathbb{P}xy)$$

is equivalent to:

$$D!x \,\&\, \exists y(Numbers(y, [\lambda z\, \mathbb{P}^{+}zx]) \,\&\, \mathbb{P}xy)$$

So it follows by GEN and RN that these formulas are necessarily and universally (for all x) equivalent. So our theorem follows from (273.13) and (49). $\bowtie$

(817.5) Since $[\lambda x\, D!x \,\&\, \mathbb{P}x\#[\lambda z\, \mathbb{P}^{+}zx]]\downarrow$, by (273.13), then from the easily-established fact, by (803.5), that:

$$\Box\forall x(D!x \,\&\, \mathbb{P}x\#[\lambda z\, \mathbb{P}^{+}zx] \equiv \mathbb{P}x\#[\lambda z\, \mathbb{P}^{+}zx]),$$

we can appeal to an instance of axiom (49) to conclude $[\lambda x\, \mathbb{P}x\#[\lambda z\, \mathbb{P}^{+}(z,x)]]\downarrow$. $\bowtie$

(817.6) We want to show:

$$\forall n \exists y(Numbers(y, [\lambda z\, \mathbb{P}^{+}zn]) \,\&\, \mathbb{P}ny)$$

We prove this by appeal to (Frege's formulation of) the principle of induction (812). So we have to find a property F such that:

- $F0$ and $\forall n \forall m(\mathbb{P}nm \rightarrow (Fn \rightarrow Fm))$, and
- the conclusion $\forall nFn$ that follows from these (by the principle of induction) is equivalent to our theorem.

Consider the property:

$$[\lambda x\, \exists y(Numbers(y, [\lambda z\, \mathbb{P}^{+}zx]) \,\&\, \mathbb{P}xy)]$$

which exists by (817.4). It is therefore a consequence of β-Conversion that:

$$[\lambda x\, \exists y(Numbers(y,[\lambda z\, \mathbb{P}^{+}zx])\,\&\, \mathbb{P}xy)]x \equiv \exists y(Numbers(y,[\lambda z\, \mathbb{P}^{+}zx])\,\&\, \mathbb{P}xy)$$

Since this equivalence holds for any object x, it holds for any natural number. So the following equivalence is a theorem:

$$[\lambda x\, \exists y(Numbers(y,[\lambda z\, \mathbb{P}^{+}zx])\,\&\, \mathbb{P}xy)]n \equiv \exists y(Numbers(y,[\lambda z\, \mathbb{P}^{+}zn])\,\&\, \mathbb{P}ny)$$

Then if we let Q be the property $[\lambda x \exists y(Numbers(y,[\lambda z \mathbb{P}^{+}zx])\&\mathbb{P}xy)]$, the claim $\forall nQn$ becomes equivalent to our theorem. So to prove our theorem by an appeal to mathematical induction (812), we have to show:

Base case: $Q0$

Inductive case: $\forall n\forall m(\mathbb{P}nm \rightarrow (Qn \rightarrow Qm))$

Base case: Show $Q0$, i.e., $[\lambda x\, \exists y(Numbers(y,[\lambda z\, \mathbb{P}^{+}zx])\,\&\, \mathbb{P}xy)]0$. Since Q exists, it suffices to show $\exists y(Numbers(y,[\lambda z\, \mathbb{P}^{+}z0])\,\&\, \mathbb{P}0y)$, by β-Conversion. Now for every G, there is an object y such that $Numbers(y,G)$ (763.1). So let a be such that $Numbers(a,[\lambda z\, \mathbb{P}^{+}z0])$. It therefore remains and suffices to show $\mathbb{P}0a$. By the main theorem governing $\mathbb{P}$ (801.3), we have to show:

$$\exists F\exists u(Fu \,\&\, Numbers(a,F) \,\&\, Numbers(0,F^{-u}))$$

where u ranges over discernible objects. Before we identify the witnesses to the above, note that 0 is a discernible object, since it is a natural number (808) and natural numbers are discernible (813.2). So we can show that the claim displayed immediately above holds if we pick our witness for F to be $[\lambda z\, \mathbb{P}^{+}z0]$ and pick our witness for u to be 0, for it then remains to show:

(ϑ) $[\lambda z\, \mathbb{P}^{+}z0]0$

(ξ) $Numbers(a,[\lambda z\, \mathbb{P}^{+}z0])$

(ζ) $Numbers(0,[\lambda z\, \mathbb{P}^{+}z0]^{-0})$

To show that (ϑ) holds, we have to show $[\lambda z\, \mathbb{P}^{+}z0]0$, i.e., by β-Conversion, that $\mathbb{P}^{+}00$. But since 0 is a discernible object, it follows by (273.30) that $0 =_D 0$, and so $\mathbb{P}^{+}00$, by the principal fact about the weak ancestral of predecessor (806.2).

(ξ) holds by assumption.

To show (ζ), i.e., $Numbers(0,[\lambda z\, \mathbb{P}^{+}z0]^{-0})$, it suffices, by (784.1), to show that:

$$\neg\exists u([\lambda z\, \mathbb{P}^{+}z0]^{-0}u)$$

For reductio, assume there is, and suppose b is such a discernible object, so that we know $[\lambda z\, \mathbb{P}^{+}z0]^{-0}b$. Since $[\lambda z\, \mathbb{P}^{+}z0]\downarrow$, $0\downarrow$, and 0 is discernible (exercises), it follows by the definition of $[\lambda z\, \mathbb{P}^{+}z0]^{-0}$ (752.2) that $[\lambda z\, [\lambda z\, \mathbb{P}^{+}z0]z \,\&\, z \neq 0]b$. So by β-Conversion, $[\lambda z\, \mathbb{P}^{+}z0]b \,\&\, b \neq 0$. And by β-Conversion and a Rule of

Substitution, $\mathbb{P}^{+}b0 \,\&\, b \neq 0$. The second conjunct implies $b \neq_D 0$, by (273.19) and (273.25), and so by the relevant instance of (806.2), it follows that $\mathbb{P}^{*}b0$, which contradicts (805.3).

Inductive Case: We want to show $\forall n \forall m(\mathbb{P}nm \rightarrow (Qn \rightarrow Qm))$, where Q is $[\lambda x\, \exists y(Numbers(y,[\lambda z\, \mathbb{P}^{+}zx]) \,\&\, \mathbb{P}xy)]$. By GEN, we have to show:

$$\mathbb{P}nm \rightarrow (Qn \rightarrow Qm)$$

By definition of Q, β-Conversion, and a Rule of Substitution, we have to show:

$$\mathbb{P}nm \rightarrow \\ \exists y(Numbers(y,[\lambda z\, \mathbb{P}^{+}zn]) \,\&\, \mathbb{P}ny) \rightarrow \exists y(Numbers(y,[\lambda z\, \mathbb{P}^{+}zm]) \,\&\, \mathbb{P}my)$$

So assume:

(A) $\mathbb{P}nm$

(B) $\exists y(Numbers(y,[\lambda z\, \mathbb{P}^{+}zn]) \,\&\, \mathbb{P}ny)$

Suppose b is the witness to (B), so that we know $Numbers(b,[\lambda z \mathbb{P}^{+}zn])$ and $\mathbb{P}nb$. Now we need to find a witness to $\exists y(Numbers(y,[\lambda z\, \mathbb{P}^{+}zm]) \,\&\, \mathbb{P}my)$. Again, since for every G, there is an object y such that $Numbers(y,G)$ (763.1), let c be such that $Numbers(c,[\lambda z \mathbb{P}^{+}zm])$. It therefore remains and suffices to show $\mathbb{P}mc$. By the principal fact about Predecessor (801.3), we have to show:

(C) $\exists F \exists u(Fu \,\&\, Numbers(c,F) \,\&\, Numbers(m,F^{-u}))$

To establish (C), we choose $[\lambda z \mathbb{P}^{+}zm]$ as our witness for F, and m as the witness for u. (Note that since m is a natural number, it is a discernible object (813.2) and so can serve as a witness for $\exists u$.) Then we have to show:

(ϑ) $[\lambda z\, \mathbb{P}^{+}zm]m$

(ξ) $Numbers(c,[\lambda z\, \mathbb{P}^{+}zm])$

(ζ) $Numbers(m,[\lambda z\, \mathbb{P}^{+}zm]^{-m})$

Show (ϑ). Since $[\lambda z\, \mathbb{P}^{+}zm]\!\downarrow$ (exercise), it suffices, by β-Conversion, to show $\mathbb{P}^{+}mm$. But since m is a natural number, it is discernible (813.2). So $m =_D m$, by (273.19). Hence $\mathbb{P}^{+}mm$, by the principle fact about the weak ancestral of predecessor (806.2).

Show (ξ). This holds by assumption.

Show (ζ). We start with the fact that the following is an instance of lemma (817.2):

$$(\mathbb{N}m \,\&\, \mathbb{P}nm) \rightarrow (Numbers(m,[\lambda z\, \mathbb{P}^{+}zn]) \equiv Numbers(m,[\lambda z\, \mathbb{P}^{+}zm]^{-m}))$$

Since $\mathbb{N}m$ (by hypothesis) and $\mathbb{P}nm$ (by assumption), it follows that:

(D) $Numbers(m,[\lambda z\,\mathbb{P}^{+}zn]) \equiv Numbers(m,[\lambda z\,\mathbb{P}^{+}zm]^{-m})$

Note that $\mathbb{P}nm$, by (A), and $\mathbb{P}nb$, by hypothesis. So $m=b$, by the functionality of predecessor (816). Since we also know $Numbers(b,[\lambda z\,\mathbb{P}^{+}zn])$ by hypothesis, it follows that $Numbers(m,[\lambda z\,\mathbb{P}^{+}zn])$. So by (D), $Numbers(m,[\lambda z\,\mathbb{P}^{+}zm]^{-m})$. ⋈

(817.7) Let n be an arbitrary natural number, so that by ∀I, it suffices show $\mathbb{P}n\#[\lambda z\mathbb{P}^{+}zn]$. Then by (817.6), it follows that $\exists y(Numbers(y,[\lambda z\mathbb{P}^{+}zn])\,\&\,\mathbb{P}ny)$. Suppose a is such an object, so that we know $Numbers(a,[\lambda z\mathbb{P}^{+}zn])\,\&\,\mathbb{P}na$. Given the second conjunct, it suffices only to show $a=\#[\lambda z\,\mathbb{P}^{+}zn]$. Since $\mathbb{P}^{+}$ is rigid (806.3), it is easy to establish that $Rigid([\lambda z\,\mathbb{P}^{+}zy])$ (exercise). So by (774.5):

(ϑ) $Numbers(\#[\lambda z\,\mathbb{P}^{+}zy],[\lambda z\,\mathbb{P}^{+}zy])$

But, independently, we know $\exists!xNumbers(x,[\lambda z\,\mathbb{P}^{+}zy])$, since (763.2) holds for every G. So (ϑ) and the assumption that $Numbers(a,[\lambda z\,\mathbb{P}^{+}zn])$ imply that $a=\#[\lambda z\,\mathbb{P}^{+}zn]$. ⋈

(817.7) [This proof more closely follows Frege's original, modulo the facts that we are working within a modal framework, making use an $\mathcal{A}$ operator, and keeping the focus on discernible objects. Nevertheless, in the proof by induction below, the base case corresponds to Frege's Theorem 154 (1893 [2013, 147]), while the inductive case corresponds to Frege's Theorem 155 (1893 [2013, 149]).] We prove this by appeal to (Frege's formulation of) the principle of induction (812). So we have to find a property F such that:

- $F0$ and $\forall n\forall m(\mathbb{P}nm \to (Fn \to Fm))$, and
- the conclusion $\forall nFn$ that follows from these by the principle of induction is equivalent to our theorem $\forall n\mathbb{P}n\#[\lambda z\,\mathbb{P}^{+}zn]$.

Consider the property:

$$[\lambda x\,\mathbb{P}x\#[\lambda z\,\mathbb{P}^{+}zx]]$$

which exists by (817.5). It is therefore a consequence of β-Conversion that:

$$[\lambda x\,\mathbb{P}x\#[\lambda z\,\mathbb{P}^{+}zx]]y \equiv \mathbb{P}y\#[\lambda z\,\mathbb{P}^{+}zy]$$

Since this equivalence holds for any object y, it holds for any natural number. So that the following equivalence is a theorem:

$$[\lambda x\,\mathbb{P}x\#[\lambda z\,\mathbb{P}^{+}zx]]n \equiv \mathbb{P}n\#[\lambda z\,\mathbb{P}^{+}zn]$$

Then if we let Q be the property $[\lambda x\,\mathbb{P}x\#[\lambda z\,\mathbb{P}^{+}zx]]$, the claim $\forall nQn$ becomes equivalent to our theorem $\forall n\mathbb{P}n\#[\lambda z\,\mathbb{P}^{+}zn]$. So to prove our theorem by an appeal to mathematical induction (812), we have to show:

Base case: $Q0$ (Frege 1893, Theorem 154)

Inductive case: $\forall n \forall m(\mathbb{P}nm \rightarrow (Qn \rightarrow Qm))$ (Frege 1893, Theorem 155)

Base case: Show $Q0$, i.e., $[\lambda x\, \mathbb{P}x\#[\lambda z\, \mathbb{P}^+ zx]]0$. Since Q exists, it suffices to show $\mathbb{P}0\#[\lambda z\, \mathbb{P}^+ z0]$, by β-Conversion. By the main theorem governing $\mathbb{P}$ (801.3), we have to show:

$$\exists F \exists u(Fu \;\&\; \#[\lambda z\, \mathbb{P}^+ z0] = \#F \;\&\; 0 = \#F^{-u})$$

where u ranges over discernible objects. Before we identify the witnesses to the above, note that 0 is a discernible object, since it is a natural number (808) and natural numbers are discernible (813.2). So if we pick our witness for F to be $[\lambda z\, \mathbb{P}^+ z0]$ and pick our witness for u to be 0, then it remains to show:

(ϑ) $[\lambda z\, \mathbb{P}^+ z0]0$

(ξ) $\#[\lambda z\, \mathbb{P}^+ z0] = \#[\lambda z\, \mathbb{P}^+ z0]$

(ζ) $0 = \#[\lambda z\, \mathbb{P}^+ z0]^{-0}$

To show that (ϑ) holds, we have to show $[\lambda z\, \mathbb{P}^+(z,0)]0$, i.e., by β-Conversion, that $\mathbb{P}^+(0,0)$. But since 0 is a discernible object, it follows by (273.30) that $0 =_D 0$, and so $\mathbb{P}^+(0,0)$, by the principal fact about the weak ancestral of predecssor (806.2).

To show that (ξ) holds, we need only note that (a) $[\lambda z\, \mathbb{P}^+ z0]\downarrow$ (exercise), and hence $\#[\lambda z\, \mathbb{P}^+ z0]\downarrow$ (exercise), and (b) everything is self-identical.

To show (ζ) holds, it suffices, by (784.3), to show:

$$\neg \exists v \mathcal{A}[\lambda z\, \mathbb{P}^+ z0]^{-0} v$$

For reductio, assume that there is such a discernible object, say b, so that we know $\mathcal{A}[\lambda z\, \mathbb{P}^+ z0]^{-0} b$. Since $[\lambda z\, \mathbb{P}^+ z0]\downarrow$, $0\downarrow$, and 0 is a discernible object, we can apply the definition of $[\lambda z\, \mathbb{P}^+ z0]^{-0}$ (752.2) to conclude:

$$\mathcal{A}[\lambda z\, [\lambda z\, \mathbb{P}^+ z0]z \;\&\; z \neq 0]b$$

Since $[\lambda z\, [\lambda z\, \mathbb{P}^+ z0]z \;\&\; z \neq 0]\downarrow$ (exercise), it follows by β-Conversion and a Rule of Substitution that: $\mathcal{A}([\lambda z\, \mathbb{P}^+ z0]b \;\&\; b \neq 0])$. Again, by β-Conversion and a Rule of Substitution, it follows that $\mathcal{A}(\mathbb{P}^+ b0 \;\&\; b \neq 0)$. Hence, by (139.2), we know both:

(A) $\mathcal{A}\mathbb{P}^+ b0$

(B) $\mathcal{A}b \neq 0$

Note that (A) implies $\mathcal{A}\mathbb{P}^* b0 \vee \mathcal{A}b =_D 0$:

Proof. By a relevant instance of (806.2) and a Rule of Substitution, (A) implies $\mathcal{A}(\mathbb{P}^*b0 \vee b =_D 0)$. But $\mathcal{A}$ distributes over a disjunction (139.9), and so it follows that $\mathcal{A}\mathbb{P}^*b0 \vee \mathcal{A}b =_D 0$.

Note also that (B) implies $\neg\mathcal{A}b =_D 0$:

The contrapositive of theorem (273.19) is $\neg b = 0 \rightarrow \neg b =_D 0$. So by Rule RA, $\mathcal{A}(\neg b = 0 \rightarrow \neg b =_D 0)$. Now (B) implies $\mathcal{A}\neg b = 0$, by (24) and a Rule of Substitution. So by (131), i.e., $\mathcal{A}(\varphi \rightarrow \psi) \rightarrow (\mathcal{A}\varphi \rightarrow \mathcal{A}\psi)$, we may infer, from what we have established, that $\mathcal{A}\neg(b =_D 0)$. So $\neg\mathcal{A}(b =_D 0)$, by (44.1).

From the two consequences of (A) and (B) just noted, it follows that $\mathcal{A}\mathbb{P}^*b0$. Hence, $\exists x\mathcal{A}\mathbb{P}^*x0$, which implies $\mathcal{A}\exists x\mathbb{P}^*x0$ (139.10). But by Rule RA, the actualization of theorem (805.3) is also a theorem, and so we know $\mathcal{A}\neg\exists x\mathbb{P}^*x0$. So by axiom (44.2), this last fact implies $\neg\mathcal{A}\exists x\mathbb{P}^*x0$. Contradiction.

Inductive Case: Show $\forall n\forall m(\mathbb{P}nm \rightarrow (Qn \rightarrow Qm))$, where Q is $[\lambda x\, \mathbb{P}x\#[\lambda z\mathbb{P}^+zx]]$. The proof of this claim appeals to (817.3), which tells us that if y precedes a natural number x, then the number of *being a weak predecessor ancestor of* y is identical to the number of *being a weak predecessor ancestors of* x *but not* x. Now by GEN, we have to show:

$$\mathbb{P}nm \rightarrow (Qn \rightarrow Qm)$$

By definition of Q, β-Conversion, and a Rule of Substitution, we have to show:

$$\mathbb{P}nm \rightarrow (\mathbb{P}n\#[\lambda z\, \mathbb{P}^+zn]) \rightarrow \mathbb{P}m\#[\lambda z\, \mathbb{P}^+zm])$$

So assume:

(C) $\mathbb{P}nm$

(D) $\mathbb{P}n\#[\lambda z\, \mathbb{P}^+zn]$

to show $\mathbb{P}m\#[\lambda z\, \mathbb{P}^+zm]$. By the principal fact about Predecessor (801.3), we have to show:

$$\exists F\exists u(Fu \,\&\, \#[\lambda z\, \mathbb{P}^+zm] = \#F \,\&\, m = \#F^{-u})$$

Note that since m is a natural number, it is a discernible object (813.2). So we can prove that the above holds if we take our witness for F to be $[\lambda z\, \mathbb{P}^+zm]$ and take our witness for u to be m. Then we have to show:

(ϑ) $[\lambda z\, \mathbb{P}^+zm]m$

(ξ) $\#[\lambda z\, \mathbb{P}^+zm] = \#[\lambda z\, \mathbb{P}^+zm]$

(ζ) $m = \#[\lambda z\, \mathbb{P}^+zm]^{-m}$

Show (ϑ). Since $[\lambda z\, \mathbb{P}^{+}zm]\downarrow$ (exercise), it suffices, by β-Conversion, to show $\mathbb{P}^{+}mm$. But since m is a natural number, it is discernible (813.2). So $m =_D m$, by (273.19). Hence $\mathbb{P}^{+}mm$, by the principle fact about the weak ancestral of predecessor (806.2).

Show (ξ). This holds by the fact that $\#[\lambda z\, \mathbb{P}^{+}zm]\downarrow$ (exercise) and the logic of identity.

Show (ζ), i.e., $m = \#[\lambda z\, \mathbb{P}^{+}zm]^{-m}$. Recall our assumptions (C) and (D), i.e., $\mathbb{P}nm$ and $\mathbb{P}n\#[\lambda z\, \mathbb{P}^{+}zn]$. It follows by the functionality of Predecessor (802.4) that $m = \#[\lambda z\, \mathbb{P}^{+}zn]$. So, by Rule =E, ($\zeta$) follows immediately from:

$$\#[\lambda z\, \mathbb{P}^{+}zn] = \#[\lambda z\, \mathbb{P}^{+}zm]^{-m}$$

But this, in turn, is a consequence of lemma (817.3) once we substitute m for x and n for y, since we know $\mathbb{N}m$, by hypothesis, and $\mathbb{P}nm$, by (C). ⋈

(818) By GEN, it suffices to show $\exists! m \mathbb{P}nm$. But since predecessor is a functional relation (816), it suffices to show that $\exists m \mathbb{P}nm$.[481] Moreover, we know that if n immediately precedes anything, that thing is a natural number (814.1), and so it suffices to show that $\exists y \mathbb{P}ny$. But this follows, *a fortiori* from (817.6), which tells us $\exists y(Numbers(y, [\lambda z\, \mathbb{P}^{+}zn])\, \&\, \mathbb{P}ny)$. ⋈

(818) [A Frege-style proof that makes use of (817.6).] By GEN, it suffices to show $\exists! m \mathbb{P}nm$. But since predecessor is a functional relation (816), it suffices to show that $\exists m \mathbb{P}nm$.[482] Moreover, we know that if n immediately precedes anything, that thing is a natural number (814.1), and so it suffices to show that $\exists y \mathbb{P}ny$. But this follows *a fortiori* from (817.7), which tells us that $\mathbb{P}n\#[\lambda z\, \mathbb{P}^{+}zn]$. ⋈

(822) From (818), we know $\exists! m \mathbb{P}nm$, i.e.,

$$\exists! x(\mathbb{N}x \,\&\, \mathbb{P}nx)$$

Hence by Rule RA,

$$\mathcal{A}\exists! x(\mathbb{N}x \,\&\, \mathbb{P}nx)$$

and so by (176.2):

$$(\vartheta) \quad \iota x(\mathbb{N}x \,\&\, \mathbb{P}nx)\downarrow$$

[481] For suppose $\exists m \mathbb{P}nm$ and let m_1 be such a natural number, so that we know $\mathbb{P}nm_1$. To show m_1 is unique, we have to show $\forall k(\mathbb{P}nk \to k = m_1)$. By GEN, it suffices to show $\mathbb{P}nk \to k = m_1$. So assume $\mathbb{P}nk$. Then given both $\mathbb{P}nk$ and $\mathbb{P}nm_1$ it follows by (816) that $k = m_1$.

[482] For suppose $\exists m \mathbb{P}nm$ and let m_1 be such a natural number, so that we know $\mathbb{P}nm_1$. To show m_1 is unique, we have to show $\forall k(\mathbb{P}nk \to k = m_1)$. By GEN, it suffices to show $\mathbb{P}nk \to k = m_1$. So assume $\mathbb{P}nk$. Then given both $\mathbb{P}nk$ and $\mathbb{P}nm_1$ it follows by (816) that $k = m_1$.

Now assume $n = m$. Then, $\Box n = m$, since the necessity of identity (125) holds universally. From this and the necessitation of the axiom for the substitution of identicals, it follows, by the K axiom, that:

$$\Box((\mathbb{N}x \,\&\, \mathbb{P}nx) \to (\mathbb{N}x \,\&\, \mathbb{P}mx))$$

$$\Box((\mathbb{N}x \,\&\, \mathbb{P}mx) \to (\mathbb{N}x \,\&\, \mathbb{P}nx)).$$

Hence by (158.4):

$$\Box((\mathbb{N}x \,\&\, \mathbb{P}nx) \equiv (\mathbb{N}x \,\&\, \mathbb{P}mx))$$

Since this holds for any x, it follows by GEN that:

$$\forall x \Box((\mathbb{N}x \,\&\, \mathbb{P}nx) \equiv (\mathbb{N}x \,\&\, \mathbb{P}mx))$$

So by the Barcan Formula (167.1):

$$(\xi)\quad \Box\forall x((\mathbb{N}x \,\&\, \mathbb{P}nx) \equiv (\mathbb{N}x \,\&\, \mathbb{P}mx))$$

Hence from (ϑ) and (ξ) it follows by (149.3) that:

$$\imath x(\mathbb{N}x \,\&\, \mathbb{P}nx) = \imath x(\mathbb{N}x \,\&\, \mathbb{P}mx)$$

If we use k as a restricted variable ranging over numbers, the above becomes:

$$\imath k \mathbb{P}nk = \imath k \mathbb{P}mk$$

So by definition (821), $n' = m'$. ⋈

(823) We prove this by mathematical induction (812). We run the induction over the property:

$$[\lambda m\, m =_D 0 \vee \exists n(m =_D \imath m P nm)]$$

Call this property S. Since S exists (exercise), it follows by β-Conversion that:

$$(\vartheta)\quad Sm \equiv m =_D 0 \vee \exists n(m =_D \imath m P nm)$$

For induction we need only show $S0 \,\&\, \mathbb{P}nm \to (Sn \to Sm)$ and we are done. It is easy to see that the left conjunct $S0$ holds, given that $0 =_D 0$ and (ϑ). For the right conjunct, assume $\mathbb{P}nm$ and Sn. It remains to show Sm. So by $\vee$I and (ϑ), it suffices to show: $\exists n(m =_D \imath m P nm)$. By assumption we have $\mathbb{P}nm$. So it suffices to show $m =_D \imath m P nm$. And by Hintikka's schema, we have to show:

$$\mathcal{A}\mathbb{P}nm \,\&\, \forall j(\mathcal{A}\mathbb{P}jm \to j = n)$$

It follows from our assumption, by (802.1), that $\Box\mathbb{P}nm$. And by (132), we have the first conjunct, $\mathcal{A}\mathbb{P}nm$.

To show the second conjunct, it suffices by GEN to show $\mathcal{A}\mathbb{P}jm \to j = n$. So assume $\mathcal{A}\mathbb{P}jm$. Note independently that if we apply RN to relevant instances of the modally strict theorem (802.1), we know $\Box(\mathbb{P}jm \to \Box\mathbb{P}jm)$. So by (174.2), $\mathcal{A}\mathbb{P}jm \equiv \mathbb{P}jm$. Hence $\mathbb{P}jm$. So it follows by the fact that $\mathbb{P}$ is a one-to-one relation (802.3)) and the definition of one-to-one relation (796.1), that $j = n$. ⋈

(824) (Exercise) [Hint: The key theorem is an instance of (153.2), but to make proper use of this instance, one needs to appeal to the rigidity of $\mathbb{P}$ (802.2), the uniqueness of successor (818), and the description that defines successor notation (821).]

(826.n) (Exercises)

(827.n) (Exercises)

(829) (Exercise) [Hint: The easiest proof may be by reductio.]

(831.1) (→) We may reason in this direction as follows:

$$\begin{array}{rcll} x < y & \equiv & \mathbb{P}^*_{\restriction\mathbb{N}}xy & \text{by definition (830.1)} \\ & \equiv & [\lambda xy\, \mathbb{N}x \,\&\, \mathbb{P}^*xy]xy & \text{by definition (828.1)} \\ & \equiv & \mathbb{N}x \,\&\, \mathbb{P}^*xy & \text{by } \beta\text{-Conversion, given } [\lambda xy\, \mathbb{N}x \,\&\, \mathbb{P}^*xy]\downarrow \\ & \to & \mathbb{N}y & \text{by theorem (814.2)} \end{array}$$

(←) Assume $\mathbb{N}x$, $\mathbb{N}y$, and $\mathbb{P}^*xy$. Conjoining the first and third, we have $\mathbb{N}x \,\&\, \mathbb{P}^*xy$. It follows by β-Conversion that $[\lambda xy\, \mathbb{N}x \,\&\, \mathbb{P}^*xy]xy$, given that the λ-expression is significant. By definition (828.1), $\mathbb{P}^*_{\restriction\mathbb{N}}xy$. Hence, by (830.1), $x < y$. ⋈

(831.2) (→) We may reason in this direction as follows:

$$\begin{array}{rcll} x \leq y & \equiv & \mathbb{P}^+_{\restriction\mathbb{N}}xy & \text{by definition (830.2)} \\ & \equiv & [\lambda xy\, \mathbb{N}x \,\&\, \mathbb{P}^+xy]xy & \text{by definition (828.1)} \\ & \equiv & \mathbb{N}x \,\&\, \mathbb{P}^+xy & \text{by } \beta\text{-Conversion, given } [\lambda xy\, \mathbb{N}x \,\&\, \mathbb{P}^+xy]\downarrow \\ & \to & \mathbb{N}y & \text{by theorem (814.3)} \end{array}$$

(←) Assume $\mathbb{N}x$, $\mathbb{N}y$, and $\mathbb{P}^+xy$. Conjoining the first and third, we have $\mathbb{N}x \,\&\, \mathbb{P}^+xy$. It follows by β-Conversion that $[\lambda xy\, \mathbb{N}x \,\&\, \mathbb{P}^+xy]xy$, given that the λ-expression is significant. By definition (828.1), $\mathbb{P}^+_{\restriction\mathbb{N}}xy$. Hence, by (830.2), $x \leq y$. ⋈

(832.1) – **(832.4)** (Exercises)

(833.1) Assume $\mathbb{P}nm$. Then by a fact about $\mathbb{P}^*$ (789.1), it follows that $\mathbb{P}^*nm$. So by (832.1), $n < m$. ⋈

(833.2) (Exercise)

(833.3) Assume $n < m$. Then by (832.1), $\mathbb{P}^*nm$. So by ∨I, $\mathbb{P}^*nm \vee n =_D m$. Hence $\mathbb{P}^+nm$, by theorem (806.2). So by (832.2), $n \leq m$. ⋈

(833.4) Assume $n \leq m$ and $n \neq m$. From the former, it follows by theorem (832.2) that $\mathbb{P}^{+}nm$. Now, we know, as an instance of theorem (273.19), that $n =_D m \rightarrow n = m$. Hence, from $n \neq m$, it follows that $\neg(n =_D m)$. So by theorem (806.2), $\mathbb{P}^{*}nm$. And by theorem (832.1), $n < m$. ⋈

(833.5) Assume $n < m$ and $m < k$, so that we know, by theorem (832.1) that $\mathbb{P}^{*}nm$ and $\mathbb{P}^{*}mk$. But $\mathbb{P}$ is a relation and so by the transitivity of its strong ancestral $\mathbb{P}^{*}$ (789.6), it follows that $\mathbb{P}^{*}nk$. So by (832.1), $n < k$. ⋈

(833.6) Assume $n \leq m$ and $m \leq k$. Then by (832.2), we know both $\mathbb{P}^{+}nm$ and $\mathbb{P}^{+}mk$. But $\mathbb{P}$ is a rigid one-to-one relation and so its weak ancestral $\mathbb{P}^{+}$ is transitive (795.6). Hence $\mathbb{P}^{+}nk$, and so by (832.2), $n \leq k$. ⋈

(833.7) Assume $n < m$ and $m \leq k$. We reason by cases. If $m = k$, then $n < k$. If $m \neq k$, then by (833.4), $m < k$. So by transitivity of $<$ (833.5), $n < k$. ⋈

(833.8) (Exercise)

(834.1) By theorem (824), we know $\mathbb{P}nn'$. So by theorem (789.1), $\mathbb{P}^{*}nn'$. Hence, by (832.1), $n < n'$.

(834.2) (Exercise)

(834.3) Assume $1 \leq n$. We also know $0 < 0'$, by (834.1). It follows from this last fact by definition of One (825.1), that $0 < 1$. So by (833.7), $0 < n$. ⋈

(834.4) Assume $m' \leq n$. By theorem (824), we know $\mathbb{P}mm'$. Hence by (833.1), $m < m'$. But from this and our assumption, it follows by (833.7), that $m < n$. ⋈

(834.5) (Exercise)

(835) We prove this by induction on m. **Base Case**: $m = 0$. Assume $Numbers(n,F)$, $Numbers(m,G)$, and $\forall u(Fu \rightarrow Gu)$. Then $Numbers(0,G)$, and so by (784.1), $\neg\exists uGu$. Since $\forall u(Fu \rightarrow Gu)$ it follows that $\neg\exists uFu$. Hence again by (784.1), $Numbers(0,F)$. Since a unique object numbers the property F (763.2), $n = 0$. So since $0 \leq 0$ (833.2), it follows that $n \leq m$.

Inductive Case: Assume that our theorem holds when $m = k$. Then our inductive hypothesis is:

(IH) $(Numbers(n,F) \,\&\, Numbers(k,G) \,\&\, \forall u(Fu \rightarrow Gu)) \rightarrow n \leq k$

and we need to show:

$$(Numbers(n,F) \,\&\, Numbers(k',G) \,\&\, \forall u(Fu \rightarrow Gu)) \rightarrow n \leq k'$$

So assume $Numbers(n,F)$, $Numbers(k',G)$, and $\forall u(Fu \rightarrow Gu)$. We reason by cases from $\forall u(Fu \equiv Gu)$ or $\neg\forall u(Fu \equiv Gu)$. If $\forall u(Fu \equiv Gu)$, then by (766), $Numbers(k',F)$. Hence, by (763.2), $n = k'$. Since $k' \leq k'$ (833.2), $n \leq k'$. If $\neg\forall u(Fu \equiv Gu)$, then given our initial assumption that $\forall u(Fu \rightarrow Gu)$, it must be that $\exists u(Gu \,\&\, \neg Fu)$. Let a be the (discernible) witness, so that we know $Ga \,\&\, \neg Fa$. Since $\forall u(Fu \rightarrow Gu)$ and $\neg Fa$, it follows that $\forall u(Fu \rightarrow (Gu \,\&\, u \neq a))$. Hence $\forall u(Fu \rightarrow G^{-a}u)$, by definition of G^{-a} (752.2), β-Conversion and a Rule of Substitution. Now let b be such that $Numbers(b, G^{-a})$ (763.1). Then by (801.3), it follows that $\mathbb{P}bk'$. By (802.3), $b = k$. Hence, $Numbers(k, G^{-a})$. Then, instantiating our IH to F and G^{-a}, it follows that $n \leq k$, and hence $n \leq k'$ by (833.6). ⋈

(836.n) (Exercises)

(838.1) Since we know $Rigid(\mathbb{P})$ (802.2) and *1-1*$(\mathbb{P})$ (802.3), we have as a consequence of (796.4):

$$\neg\mathbb{P}^*00 \rightarrow (\mathbb{P}^+0n \rightarrow \neg\mathbb{P}^*nn)$$

By theorem (805.3), we know $\neg\mathbb{P}^*00$. Hence:

(ϑ) $\mathbb{P}^+0n \rightarrow \neg\mathbb{P}^*nn$

But n is, by hypothesis, a natural number. So by (807.3), $\mathbb{P}^+0n$. So by (ϑ), $\neg\mathbb{P}^*nn$. Consequently, by (833.1), we have $\neg(n < n)$. ⋈

(838.2) By (838.1), $n \nless n$. So by (833.1), $\neg\mathbb{P}^*nn$. So by (789.1), $\neg\mathbb{P}nn$. ⋈

(838.3) By (824), we know $\mathbb{P}nn'$. For reductio, suppose $n = n'$. Then $\mathbb{P}nn$, which contradicts (838.2). ⋈

(839.2) $(\rightarrow)$ We may reason in this direction as follows, given that the λ-expression is significant:

$$\begin{array}{lcll}
x \doteq y & \equiv & x =_{D \upharpoonright \mathbb{N}} y & \text{by definition (839.1)} \\
& \equiv & [\lambda xy\, \mathbb{N}x \,\&\, x =_D y]xy & \text{by definition (828.1)} \\
& \equiv & \mathbb{N}x \,\&\, x =_D y & \text{by } \beta\text{-Conversion} \\
& \rightarrow & \mathbb{N}y & \text{(273.19) and Rule =E}
\end{array}$$

$(\leftarrow)$ Assume $\mathbb{N}x$, $\mathbb{N}y$, and $x =_D y$. Conjoining the first and third, we have $\mathbb{N}x \,\&\, x =_D y$. Since $[\lambda xy\, \mathbb{N}x \,\&\, x =_D y]$ exists, it follows by β-Conversion that $[\lambda xy\, \mathbb{N}x \,\&\, x =_D y]xy$. By definition (828.1), $x =_{D \upharpoonright \mathbb{N}} y$. Hence, by (839.1), $x \doteq y$. ⋈

(840.1) Assume $x \doteq y$. Then by (839.2), we know $\mathbb{N}x$, $\mathbb{N}y$, and $x =_D y$. But by an instance of (273.19), the last of these consequences implies $x = y$. ⋈

(840.2) Assume $\mathbb{N}x \vee \mathbb{N}y$. $(\rightarrow)$ Then $x \doteq y \rightarrow x = y$ is already theorem (840.1). $(\leftarrow)$ Assume $x = y$. To show $x \doteq y$, we have to show, by (839.2):

$$\mathbb{N}x \,\&\, \mathbb{N}y \,\&\, x =_D y$$

To do this we reason by cases from our initial assumption $\mathbb{N}x \vee \mathbb{N}y$:

Suppose $\mathbb{N}x$. Since we've assumed $x = y$, it follows that $\mathbb{N}y$. So it remains to show $x =_D y$. Now $\mathbb{N}x$ and $\mathbb{N}y$ imply, respectively, that $D!x$ and $D!y$, by (813.2). Then $x =_D y$ by (273.18).

Suppose $\mathbb{N}y$. Then our conclusion follows by analogous reasoning. ⋈

(840.3) – **(840.6)** (Exercises)

(840.7) We can reason as follows:

$$\begin{array}{llll} n \leq m & \equiv & P^{+}nm & (832.2) \\ & \equiv & P^{*}nm \vee n =_D m & (806.2) \\ & \equiv & n < m \vee n =_D m & (832.1) \\ & \equiv & n < m \vee n = m & (839.2), (840.2) \bowtie \end{array}$$

(841.2) (Exercise)

(842.1) By (836.1), $0 < 1$. Hence, $1 > 0$, by λ-Conversion and definition (830.3). So, $\mathbb{N}_{+}1$, by (841.2). ⋈

(842.2) Assume $1 \leq n$. Then by (834.3), $0 < n$, i.e., $n > 0$. Hence $\mathbb{N}_{+}1$, by (841.2). ⋈

(843) Assume $\mathbb{N}_{+}x$. Since $\mathbb{P}$ is one-to-one (802.3), it suffices to show $\exists n \mathbb{P}nx$. Our assumption implies $x > 0$, by (841.1), i.e., $0 < x$. So by (832.1), $\mathbb{P}^{*}0x$. So by (789.5), $\exists z \mathbb{P}zx$. Let a be such an object, so that we know $\mathbb{P}ax$. Since $\mathbb{P}ax$ and our assumption implies $\mathbb{N}x$ (exercise), it follows from (815) that $\mathbb{N}a$. Hence, $\exists n \mathbb{P}nx$. ⋈

(844.1) Assume $Fu \,\&\, Numbers(n, F^{-u})$. Independently, we know $\mathbb{P}(n, n')$, by (824). So by theorem (801.3), $\exists H \exists u(Hu \,\&\, Numbers(n', H) \,\&\, Numbers(n, H^{-u}))$. So let Q and b be such a property and discernible object, so that we know $Qb \,\&\, Numbers(n', Q) \,\&\, Numbers(n, Q^{-b})$. Since $Numbers(n, F^{-u})$ and $Numbers(n, Q^{-b})$, it follows that $F^{-u} \approx_D Q^{-b}$, by (764.2). Hence, by Lemma (754), $F \approx_D Q$. So since $Numbers(n', Q)$, it follow by (764.1) that $Numbers(n', F)$. ⋈

(844.2) Assume $Numbers(n', F)$. We want to show $\exists u(Fu \,\&\, Numbers(n, F^{-u}))$. Now, independently, we know $\mathbb{P}(n, n')$, by (824). So by theorem (801.3):

$$\exists H \exists u(Hu \,\&\, Numbers(n', H) \,\&\, Numbers(n, H^{-u}))$$

Let Q and b be such a property and discernible object, so that we know:

$$Qb \,\&\, Numbers(n', Q) \,\&\, Numbers(n, Q^{-b})$$

Since $Numbers(n', F)$ and $Numbers(n', Q)$, it follows by (764.2) that $F \approx_D Q$ and, by symmetry, $Q \approx_D F$. Hence, by definition (747.3), there is a relation G that witnesses this equinumerosity, i.e.,

$$\exists G(G|{:}\, Q \xleftrightarrow{1\text{-}1}_D F)$$

Let R be such a relation. Then, *a fortiori*, by (747.2), $\forall u(Qu \to \exists! v(Fv \,\&\, Ruv))$. So, in particular, $Qb \to \exists! v(Fv \,\&\, Rbv)$. Since Qb, let c be the unique discernible object such that Fc and Rbc. So we have now established $Q \approx_D F$, Qb, and Fc. So by (753), $Q^{-b} \approx_D F^{-c}$. But since we also have that $Numbers(n, Q^{-b})$, it follows by (764.1) that $Numbers(n, F^{-c})$. Hence we now have $Fc \,\&\, Numbers(n, F^{-c})$. So $\exists u(Fu \,\&\, Numbers(n, F^{-u}))$. ⋈

(845.2) By definition 845.1, we know $\exists_{!0} uFu \equiv Numbers(0, F)$. So, it suffices to show $Numbers(0, F) \equiv \neg\exists uFu$. But this is just (784.1). ⋈

(845.3) We want to show:

$$\exists_{!n'} uFu \equiv \exists u(Fu \,\&\, \exists_{!n} vF^{-u}v)$$

By definition (845.1), we have to show:

$$Numbers(n', F) \equiv \exists u(Fu \,\&\, \exists_{!n} vF^{-u}v)$$

($\to$) Assume $Numbers(n', F)$. We need to show $\exists u(Fu \,\&\, \exists_{!n} vF^{-u}v)$. But by definition (845.1) and a rule of substitution, it suffices to show:

$$\exists u(Fu \,\&\, Numbers(n, F^{-u}))$$

This is easy, since it follows directly from $Numbers(n', F)$ by (844.2).

($\leftarrow$) Assume $\exists u(Fu \,\&\, \exists_{!n} vF^{-u}v)$. Let a be such an object, so that we know Fa and $\exists_{!n} vF^{-a}v$. By definition (845.1), the latter implies $Numbers(n, F^{-a})$. So by (844.1), $Numbers(n', F)$. ⋈

(846.1) Since it is clear that the two individuals asserted to be identical are abstract, it suffices to show they encode the same properties, i.e., to show:

$$nG \equiv \imath x(A!x \,\&\, \forall F(xF \equiv \exists_{!n} uFu))G$$

But by (258.2), it suffices to show:

$$(\vartheta)\quad nG \equiv \mathcal{A}\exists_{!n} uGu$$

But by definition (845.1) and a Rule of Substitution (160.3), it suffices to show:

$$(\zeta)\quad nG \equiv \mathcal{A}Numbers(n, G)$$

By (813.1), we know $NaturalCardinal(n)$, so the antecedent of (780) is satisfied and it follows that:

$$(\text{A})\quad nG \equiv n = \#G$$

So we can now reason as follows:

$$\begin{array}{llll} nG & \equiv & n = \#G & \text{by (A)} \\ & \equiv & \mathit{Numbers}(n, [\lambda z\, \mathcal{A}Gz]) & \text{by (774.1)} \\ & \equiv & \mathcal{A}\mathit{Numbers}(n, G) & \text{by (769.4)} \end{array}$$

⋈

(846.2) Pick any natural number n. By (813.1), $\mathit{NaturalCardinal}(n)$. So by definition (777), there exists a property G, such that $n = \#G$. So let P be such a property, so that $n = \#P$. So if we let $[\lambda z \mathcal{A}Pz]$ be our witness, it suffices by ∃I to show $\exists_{!n} u([\lambda z\, \mathcal{A}Pz]u)$. From $n = \#P$ and (774.2), it follows that $\mathit{Numbers}(n, [\lambda z\, \mathcal{A}Pz])$. So by definition (845.1), it follows that $\exists_{!n} u([\lambda z\, Pz]u)$. ⋈

(850.1) – **(850.4)** (Exercises)

(851.1) By definition (848.1), we have to show:

$$\forall x \exists ! y [\lambda xy\, \mathcal{U}x \,\&\, y \doteq 0]xy$$

From the fact that the λ-expression is significant, β-Conversion, and a Rule of Substitution, we have to show:

$$\forall x \exists ! y (\mathcal{U}x \,\&\, y \doteq 0)$$

By GEN, we have to show:

$$\exists ! y (\mathcal{U}x \,\&\, y \doteq 0)$$

We can establish this if we choose Zero to be our witness. By the uniqueness quantifier, we have to show:

(A) $\mathcal{U}x \,\&\, 0 \doteq 0$

(B) $\forall z((\mathcal{U}x \,\&\, z \doteq 0) \rightarrow z = 0)$

(A) $\mathcal{U}x$ is an instance of (850.3). Since $\mathbb{N}0$ is a theorem (808), it follows by (840.3) that $0 \doteq 0$.

(B) By GEN, it suffices to show $\mathcal{U}x \,\&\, z \doteq 0 \rightarrow z = 0$. So assume $\mathcal{U}x$ and $z \doteq 0$. Then, again, since $\mathbb{N}0$ is a theorem, it follows from the second assumption by (840.2) that $z = 0$. ⋈

(851.2) (Exercise)

(851.3) Although the theorem is a special case of (851.2), we give a proof without assuming that theorem. Since $[\lambda y\, y \doteq 0]$ clearly exists, we have to show, by definition (848.3), β-Conversion, and a Rule of Substitution, $\exists ! y (y \doteq 0)$. If we choose Zero to be our witness, then by the definition of the uniqueness quantifier, we have to show:

(A) $0 \doteq 0$

(B) $\forall z(z \doteq 0 \rightarrow z = 0)$

(A) Since $\mathbb{N}0$ is a theorem (808), it follows by (840.3) that $0 \doteq 0$. (B) By GEN, it suffices to show $z \doteq 0 \rightarrow z = 0$. So assume $z \doteq 0$. Then since $\mathbb{N}0$ is a theorem, it follows by (840.2) that $z = 0$. ⋈

(852.1) By theorem (273.5) it is a theorem that there are discernible objects, i.e., $\exists x D!x$. So assume a is such an object, i.e., $D!a$. Then our next step is to show that the binary relation *being an x and y such that x exemplifies $\mathcal{U}$ and y is identical to a* ($[\lambda xy\, \mathcal{U}x \,\&\, y =_D a]$) is a witness to our existential claim, i.e., our next step is to establish (ξ):

(ξ) $\mathit{Function}^1([\lambda xy\, \mathcal{U}x \,\&\, y =_D a])$

Proof. By definition (848.1), we have to show:

$$\forall x \exists! y [\lambda xy\, \mathcal{U}x \,\&\, y =_D a]xy$$

where a is a discernible object. Given β-Conversion, the fact that the λ-expression is significant, and a Rule of Substitution, we have to show:

$$\forall x \exists! y(\mathcal{U}x \,\&\, y =_D a)$$

By GEN, we have to show:

$$\exists! y(\mathcal{U}x \,\&\, y =_D a)$$

So pick our witness to be a. Then by the definition of the uniqueness quantifier and $\exists$I, we have to show both:

(ϑ) $\mathcal{U}x \,\&\, a =_D a$

(ζ) $\forall z((\mathcal{U}x \,\&\, z =_D a) \rightarrow z = a)$

But for (ϑ), we know $\mathcal{U}x$ by theorem (850.3), and we know $a =_D a$ from (273.30) and the assumption that a is discernible. For (ζ), assume, by GEN, $\mathcal{U}x \,\&\, z =_D a$. Then it follows from the second conjunct that $z = a$, by (273.19). This completes the subproof of (ξ).

Note that by (39.2), $[\lambda xy \mathcal{U}x \& y =_D a]\downarrow$, since this is a core λ-expression (9.2): no variable bound by the λ occurs in encoding position (9.1) in the matrix. Hence, by Rule $\exists$I (101.2), then follows from (ξ) that:

$$\exists R\mathit{Function}^1(R)$$

So thus far, we have shown:

$$D!a \vdash \exists R\mathit{Function}^1(R)$$

Hence, by $\exists$E:

$\exists x D!x \vdash \exists R Function^1(R)$

So, by the Deduction Theorem, $\exists x D!x \to \exists R Function^1(R)$. But since $\exists x D!x$ is a theorem, $\exists R Function^1(R)$. ⋈

(852.2) (Exercise)

(853.1) – **(853.2)** (Exercises)

(853.3) The proof starts out along the same lines as the proof in (852.1). By (273.5), it is a theorem that there are discernible objects, i.e., $\exists x D!x$. So assume a is such an object, i.e., $D!a$. Then, in the proof of (852.1), we derived that the binary relation $[\lambda xy\, \mathcal{U}x \,\&\, y =_D a]$ is a total unary function. This latter claim was labeled (ξ) in the proof of (852.1), and so one of the first steps of that proof was to establish the following derivation:

$D!a \vdash Function^1([\lambda xy\, \mathcal{U}x \,\&\, y =_D a])$

Hence, by RN:

(ϑ) $\Box D!a \vdash \Box Function^1([\lambda xy\, \mathcal{U}x \,\&\, y =_D a])$

Now independently we know, by applying GEN to the left-to-right direction of theorem (180.1), we know $\forall x(D!x \to \Box D!x)$. So as an instance we have $D!a \to \Box D!a$. Thus, by (63.10):

(ξ) $D!a \vdash \Box D!a$

So from (ξ) and (ϑ), it follows by (63.8) that:

(ζ) $D!a \vdash \Box Function^1([\lambda xy\, \mathcal{U}x \,\&\, y =_D a])$

But since the λ-expression $[\lambda xy\, \mathcal{U}x \,\&\, y =_D a]$ is significant, we can establish, by now familiar reasoning, that:

(ω) $\Box Function^1([\lambda xy\, \mathcal{U}x \,\&\, y =_D a]) \vdash \exists R \Box Function^1(R)$

So from (ζ) and (ω), it also follows by (63.8) that:

$D!a \vdash \exists R \Box Function^1(R)$

Since a was arbitrary, it follows by $\exists$E that:

$\exists x D!x \vdash \exists R \Box Function^1(R)$

But it follows from this last fact and the fact that $\exists x D!x$ is a theorem, by now familiar reasoning, that $\exists R \Box Function^1(R)$. ⋈

(853.4) (Exercise)

(854.1) We prove only the case for $n = 1$. By (217.1), there are contingently true propositions. Assume p_1 is such a proposition, so that we know $p_1 \,\&\, \Diamond\neg p_1$.[483] Our proof strategy is as follows. First we derive our theorem from the assumption $p_1 \,\&\, \Diamond\neg p_1$; i.e., we show:

(A) $p_1 \,\&\, \Diamond\neg p_1 \vdash \exists R(\mathit{Function}(R) \,\&\, \Diamond\neg \mathit{Function}(R))$

Then we conclude, by an application of ∃E:

(B) $\exists p(p \,\&\, \Diamond\neg p) \vdash \exists R(\mathit{Function}(R) \,\&\, \Diamond\neg \mathit{Function}(R))$

Then by an applications of the Deduction Theorem, we conclude:

(C) $\vdash \exists p(p \,\&\, \Diamond\neg p) \to (\exists R(\mathit{Function}(R) \,\&\, \Diamond\neg \mathit{Function}(R)))$

But since $\exists p(p \& \Diamond\neg p)$ is a theorem, we can conclude, by an application of (63.5), that:

(D) $\vdash \exists R(\mathit{Function}(R) \,\&\, \Diamond\neg \mathit{Function}(R))$

Since the inferences from (A) to (B), (B) to (C) and (C) to (D) are easy, it remains only to show (A). So given our assumption $p_1 \,\&\, \Diamond\neg p_1$, consider the relation f_1:

$$[\lambda xy\, p_1 \,\&\, \mathcal{U}x \,\&\, \,\&\, y \doteq 0] \tag{f_1}$$

To show $\exists R(\mathit{Function}(R) \,\&\, \Diamond\neg \mathit{Function}(R))$, it suffices by &I and ∃I to show:

(E) $\mathit{Function}(f_1)$

(F) $\Diamond\neg \mathit{Function}(f_1)$

(E) To show f_1 is a total function, we have to show (848.1):

$$\forall x \exists! y [\lambda xy\, p_1 \,\&\, \mathcal{U}x \,\&\, y \doteq 0]xy$$

Given that the λ-expression is significant, β-Conversion, and a Rule of Substitution, we have to show:

$$\forall x \exists! y (p_1 \,\&\, \mathcal{U}x \,\&\, y \doteq 0)$$

By GEN, we have to show:

$$\exists! y (p_1 \,\&\, \mathcal{U}x \,\&\, y \doteq 0)$$

So pick our witness to be 0. Then by the definition of the uniqueness quantifier and ∃I, we have to show both:

[483] We've noted on prior occasions why reasoning of the kind used in the remainder of this proof is modally strict notwithstanding the appeal to an assumption that isn't necessarily true. In this particular case, the assumption plays a role in *two* places in the proof. In what follows, note that (a) RN is correctly applied, and (b) the assumption is ultimately discharged in both places. See again the discussion in Remark (70), Remark (218), and the proof of theorem (221.1).

(ϑ) $p_1 \,\&\, \mathcal{U}x \,\&\, 0 \doteq 0$

(ζ) $\forall z((p_1 \,\&\, \mathcal{U}x \,\&\, z \doteq 0) \rightarrow z = 0)$

To show (ϑ), we note: p_1 is true by hypothesis; $\mathcal{U}x$ is a theorem, by (850.3); and since $\mathbb{N}0$ is a theorem (808), $0 \doteq 0$, by (840.3). To show (ζ), assume, by GEN, $p_1 \,\&\, \mathcal{U}x \,\&\, z \doteq 0$. Then it follows from the theorem $\mathbb{N}0$ and the third conjunct that $z = 0$, by (840.2).

(F) To show f_1 possibly fails to be a total function, we use the following proof strategy. We'll first prove the following conditional by modally strict means:

(G) $\neg p_1 \rightarrow \neg \mathit{Function}(f_1)$

By Rule RM$\Diamond$ (157.2), it then follows that:

(H) $\Diamond\neg p_1 \rightarrow \Diamond\neg \mathit{Function}(f_1)$

But since $\Diamond\neg p_1$ is true by hypothesis, it follows that:

(I) $\Diamond\neg \mathit{Function}(f_1)$

Now since the reasoning in the steps from (G) to (H) and from (H) to (I) is straightforward, it remains only to show (G). So assume $\neg p_1$, for conditional proof. (This assumption and the reasoning that follows is independent of our earlier assumption that $p_1 \,\&\, \Diamond\neg p_1$.) Then $\neg(p_1 \,\&\, \mathcal{U}x \,\&\, y \doteq 0)$, for if a formula χ is false, then no conjunction containing χ as a conjunct is true. Hence, by by Rule $\overrightarrow{\beta}$C (184.1.b):

$$\neg[\lambda xy\, p_1 \,\&\, \mathcal{U}x \,\&\, y \doteq 0]xy$$

I.e., $\neg f_1 xy$. Since the free occurrences of y in the above don't occur free in any hypothesis, we may use GEN to infer $\forall y \neg f_1 xy$. Hence, by quantifier negation, $\neg\exists y f_1 xy$. *A fortiori*, $\neg\exists! y f_1 xy$. So, existentially generalizing on x, we have $\exists x \neg\exists! y f_1 xy$. And again by quantifier negation, $\neg\forall x \exists! y f_1 xy$. So $\neg \mathit{Function}(f_1)$. Thus, if we discharge our assumption $\neg p_1$ by conditional proof, we have derived $\neg p_1 \rightarrow \neg \mathit{Function}(f_1)$. This conditional doesn't depend on $\neg p_1$ and doesn't depend on our assumption $p_1 \,\&\, \Diamond\neg p_1$. Since it was derived by modally strict means from no assumptions, we may apply RN and continue reasoning by way of the remaining simple steps in the proof strategy outlined above. ⋈

(854.2) (Exercise)

(857)★ ($\rightarrow$) Assume $\hat{f}(x_1, \ldots, x_n) = y$. Then by definition of $\hat{f}(x_1, \ldots, x_n)$ (856), it follows that:

$$\imath z \hat{f} x_1 \ldots x_n z = y$$

So by the symmetry of identity and (145.2)★, $\hat{f}x_1 \dots x_n y$.

($\leftarrow$) Assume $\hat{f}x_1 \dots x_n y$. Now since $\hat{f}$ is a total function, we know by definition (848.2) that $\hat{f}$ relates $x_1, \dots, x_n$ to a unique object. So it follows that $\forall z(\hat{f}x_1 \dots x_n z \rightarrow z = y)$. Thus, we have:

$$\hat{f}x_1 \dots x_n y \,\&\, \forall z(\hat{f}x_1 \dots x_n z \rightarrow z = y)$$

Hence, by the right-to-left direction of Hintikka's schema (142)★, it follows that $y = \imath y \hat{f}x_1 \dots x_n y$. Hence by definition (856), $y = \hat{f}(x_1, \dots, x_n)$. So $\hat{f}(x_1, \dots, x_n) = y$, by the symmetry of identity. ⋈

(858)★ Without loss of generality, we prove only the case for $n = 1$. Assume $\forall x \forall y(\hat{f}xy \equiv \hat{h}xy)$. Then by two applications of $\forall$E:

(ϑ) $\hat{f}xy \equiv \hat{h}xy$

We may then conclude $\hat{f}(x) = y \equiv \hat{h}(x) = y$ by the following biconditional reasoning:

$$\begin{array}{llll} \hat{f}(x) = y & \equiv & \hat{f}xy & \text{by (857)★} \\ & \equiv & \hat{h}xy & \text{by } (\vartheta) \\ & \equiv & \hat{h}(x) = y & \text{by (857)★} \end{array}$$

Since the variable y doesn't occur free in the only assumption we've made to derive $\hat{f}(x) = y \equiv \hat{h}(x) = y$, we may apply GEN to conclude:

$$\forall y(\hat{f}(x) = y \equiv \hat{h}(x) = y)$$

Hence, by the right-to-left direction of (117.4) it follows that $\hat{f}(x) = \hat{h}(x)$. By the necessity of identity (125.2), it follows that $\Box(\hat{f}(x) = \hat{h}(x))$. Now since x doesn't occur free in the only assumption we've made to derive our last result, it follows by GEN that $\forall x \Box(\hat{f}(x) = \hat{h}(x))$. Hence by the Barcan Formula (167.1), it follows that $\Box \forall x(\hat{f}(x) = \hat{h}(x))$. ⋈

(863.1) In what follows, we use w as an individual variable (not restricted to possible worlds). By definition (862.1), we have to show:

$$\forall x_1 \forall x_2([\lambda xy\, \mathcal{U}x \,\&\, D!y]x_1 x_2 \rightarrow \exists! w(D!w \,\&\, [\lambda xyz\, \mathcal{U}x \,\&\, y =_D z]x_1 x_2 w))$$

To avoid clash of variables we prove this by $\forall$I instead of by GEN. So, where a and b are arbitrary, assume $[\lambda xy\, \mathcal{U}x \,\&\, D!y]ab$, to show $\exists! w(D!w \,\&\, [\lambda xyz\, \mathcal{U}x \,\&\, y =_D z]abw)$. Since the λ-expression is significant, β-Conversion transforms our assumption into:

(ϑ) $\mathcal{U}a \,\&\, D!b$

Moreover, β-Conversion yields the modally strict theorem that:

$$[\lambda xyz\, \mathcal{U}x \,\&\, y =_D z]abw \equiv \mathcal{U}a \,\&\, b =_D w$$

So, to establish our conclusion, it suffices by a Rule of Substitution to show: $\exists!w(D!w \,\&\, \mathcal{U}a \,\&\, b =_D w)$, i.e., we have to show:

$$\exists w((D!w \,\&\, \mathcal{U}a \,\&\, b =_D w) \,\&\, \forall z((D!z \,\&\, \mathcal{U}a \,\&\, b =_D z) \to z = w))$$

But we can prove this if we choose b as our witness. We need to show:

$$D!b \,\&\, \mathcal{U}a \,\&\, b =_D b \,\&\, \forall z((D!z \,\&\, \mathcal{U}a \,\&\, b =_D z) \to z = b)$$

$D!b$ is known by (ϑ). We also know $\forall x \mathcal{U}x$, by theorem (850.3). So $\mathcal{U}a$. Moreover, since $D!b$, it follows by (273.30) that $b =_D b$. So it remains to show uniqueness. By GEN, it suffices to show $(D!z \,\&\, \mathcal{U}a \,\&\, b =_D z) \to z = b$. So assume $D!z \,\&\, \mathcal{U}a \,\&\, b =_D z$. But the third conjunct implies $b = z$ by (273.19), which in turn yields $z = b$, by the symmetry of identity. ⋈

(863.2) By definition (743.1), we have to show:

$$\forall x(\mathbb{N}x \to \exists!y(\mathbb{N}y \,\&\, x \doteq y))$$

So, by GEN, assume $\mathbb{N}x$. By definition of the uniqueness quantifier, we have to show:

$$\exists y((\mathbb{N}y \,\&\, x \doteq y) \,\&\, \forall z((\mathbb{N}z \,\&\, x \doteq z) \to z = y))$$

But we can prove this if we let x be our witness. For then we have to show:

$$\mathbb{N}x \,\&\, x \doteq x \,\&\, \forall z((\mathbb{N}z \,\&\, x \doteq z) \to z = x)$$

Now we have $\mathbb{N}x$ by assumption. So we may apply (840.3) to conclude $x \doteq x$. It then remains to show uniqueness. Assume $\mathbb{N}z \,\&\, x \doteq z$. Then by (840.2), it follows that $x = z$. So $z = x$ by symmetry of identity. ⋈

(863.3) (Exercise)

(864) Theorem (818) is $\forall n \exists! m \mathbb{P}nm$. By expanding our restricted variables, this becomes:

$$\forall x(\mathbb{N}x \to \exists!y(\mathbb{N}y \,\&\, \mathbb{P}xy))$$

Hence, by definition (743.1), $\mathbb{P}\,|\!: \mathbb{N} \to \mathbb{N}$. ⋈

(865.1) – **(865.2)** (Exercises)

(865.3) (Although this theorem is a special case of (865.2), we give an independent proof.) Assume $R\,|\!: p \longrightarrow G$. Then by (862.2):

(ϑ) $p \to \exists!y(Gy \,\&\, Ry)$

To show $R_{\upharpoonright p}\,|\!: p \longrightarrow G$, we have to show, by (862.2):

$$p \to \exists!y(Gy \,\&\, R_{\upharpoonright p}y)$$

Since $R_{\upharpoonright p}$ is defined as $[\lambda y\, p \,\&\, Ry]$ (828.3), which clearly exists, we have to show, by β-Conversion and a Rule of Substitution:

(A) $p \rightarrow \exists!y(Gy \,\&\, p \,\&\, Ry)$

So assume p. Then by (ϑ), $\exists!y(Gy \,\&\, Ry)$. Suppose a is such an object, so that by the definition of the uniqueness quantifier, we know:

(ξ) $Ga \,\&\, Ra \,\&\, \forall z(Gz \,\&\, Rz \rightarrow z = a)$

To show the consequent of (A), we choose a as our witness and show:

$$Ga \,\&\, p \,\&\, Ra \,\&\, \forall z(Gz \,\&\, p \,\&\, Rz \rightarrow z = a)$$

Since we already know Ga (ξ), p (by assumption), and Ra (ξ), it remains to show uniqueness. So assume $Gz \,\&\, p \,\&\, Rz$. Then the first and third conjuncts imply, by the third conjunct of (ξ), that $z = a$. ⋈

(867) We have to show:

(A) $[\lambda xy\, (\mathbb{N}x \,\&\, \mathbb{N}y) \rightarrow x \doteq y] \,|\!: \mathbb{N} \longrightarrow \mathbb{N}$

(B) $\forall x \forall y(\mathbb{N}x \,\&\, [\lambda xy\, (\mathbb{N}x \,\&\, \mathbb{N}y) \rightarrow x \doteq y]xy \rightarrow \mathbb{N}y)$

(A) By GEN, definition (743.1), β-Conversion, and a Rule of Substitution, we have to show:

$$\mathbb{N}x \rightarrow \exists!y(\mathbb{N}y \,\&\, ((\mathbb{N}x \,\&\, \mathbb{N}y) \rightarrow x \doteq y))$$

So assume the antecedent $\mathbb{N}x$. Then, we leave it as an exercise to show that x is a witness to the consequent.

(B) (Exercise). ⋈

(868) We have to show:

(A) $\mathbb{P} \,|\!: \mathbb{N} \longrightarrow \mathbb{N}$

(B) $\forall x \forall y(\mathbb{N}x \,\&\, \mathbb{P}xy \rightarrow \mathbb{N}y)$

(A) was established as theorem (864). For (B), assume by GEN that $\mathbb{N}x \,\&\, \mathbb{P}xy$. Then by (814.1), $\mathbb{N}y$. ⋈

(870.1) Assume $R \mathrel{\dot{\div}} F \longrightarrow G$. Then by definitions (866) and (743.1), we know:

(ϑ) $\forall x(Fx \rightarrow \exists!y(Gy \,\&\, Rxy))$

(ξ) $\forall x \forall y(Fx \,\&\, Rxy \rightarrow Gy)$

To show *FunctionalOn*(R, F), we have to show, by (869):

$$\forall x(Fx \rightarrow \forall y \forall z(Rxy \,\&\, Rxz \rightarrow y = z))$$

So, by GEN, assume Fx. And again by GEN, assume Rxy and Rxz. Then, by our three assumptions, it follows from (ξ) both that Gy and Gz. But our assumption Fx, by (ϑ), implies that x bears R to a unique G-object. Hence, $y = z$. ⋈

(870.2) (Exercise)

(870.3) (Although this theorem is a special case of (870.2), we give an independent proof.) Assume $R \dotdiv p \longrightarrow G$. Then by definitions (866) and (862.2), we know:

(ϑ) $p \to \exists! y(Gy \,\&\, Ry))$

(ξ) $\forall y(p \,\&\, Ry \to Gy)$

To show $FunctionalOn(R, p)$, we have to show, by (869):

$$p \to \forall y \forall z(Ry \,\&\, Rz \to y = z))$$

So assume p. Since y and z don't occur free in any assumptions, it suffices, by GEN, to assume Ry and Rz and show $y = z$. Then, by our three assumptions, it follows from (ξ) both that Gy and Gz. But our assumption p, by (ϑ), implies that there is a unique object that is both G and R. Hence, $y = z$. ⋈

(875.1) – **(875.5)** (Exercises)

(875.6) (Although this theorem is a special case of (875.5), we give an independent proof.) Assume $R \dotdiv p \longrightarrow G$. Then by (866.3), we know first that $R\,|\!: p \longrightarrow G$, i.e., by (862.2):

(ϑ) $p \to \exists! y(Gy \,\&\, Ry)$

and second that:

(ξ) $\forall y(p \,\&\, Ry \to Gy)$

Now to show $R_{\restriction p} : p \longrightarrow G$, we have to show, by (875.3), that:

(A) $p \to \exists! y(Gy \,\&\, R_{\restriction p} y)$

(B) $\forall y(p \,\&\, R_{\restriction p} y \to Gy)$

(C) $p \equiv \exists y R_{\restriction p} y$

(A) Assume p. Then by (ϑ), $\exists! y(Gy \,\&\, Ry)$. Hence $\exists! y(Gy \,\&\, p \,\&\, Ry)$. But by the definition of $R_{\restriction p}$ (828.3) and β-Conversion, it is a modally strict theorem that $R_{\restriction p} y \equiv p \,\&\, Ry$. Hence $\exists! y(Gy \,\&\, R_{\restriction p} y)$, by a Rule of Substitution.

(B) By GEN, we show $p \,\&\, R_{\restriction p} y \to Gy$. So assume both p and $R_{\restriction p} y$. The latter, by the definition of $R_{\restriction p}$ as $[\lambda y\, p \,\&\, Ry]$ (828.3), implies Ry *a fortiori*. Hence, by (ξ), Gy.

(C) By definition of $R{\restriction_p}$, β-Conversion and a Rule of Substitution, we need only show $p \equiv \exists y(p \,\&\, Ry)$. ($\rightarrow$) Assume p. Then by (ϑ), $\exists!y(Gy \,\&\, Ry)$. Suppose a is such an object so that we know *a fortiori* from the definition of the uniqueness quantifier that Ra. So $p \,\&\, Ra$. Hence, $\exists y(p \,\&\, Ry)$. ($\leftarrow$) Assume $\exists y(p \,\&\, Ry)$. Then it follows *a fortiori* that p. ⋈

(875)/Exercise 13 ($\rightarrow$) Assume R is a function on domain F, i.e., that $\exists G(R : F \longrightarrow G)$. Suppose P is a witness, so that we know $R : F \longrightarrow P$. Then by definition (874.1), it follows both that $R \dotdiv F \longrightarrow P$ and $HasDomain(R, F)$. So it remains to show $FunctionalOn(R, F)$. But this follows from $R \dotdiv F \longrightarrow P$ by (870.1). ($\leftarrow$) Assume $FunctionalOn(R, F)$ and $HasDomain(R, F)$. It follows from these two assumptions, by (869.1) and (872.1), respectively, that:

(ϑ) $\forall x(Fx \rightarrow \forall y \forall z(Rxy \,\&\, Rxz \rightarrow y = z))$

(ξ) $\forall x(Fx \equiv \exists y Rxy)$

Now to show $\exists G(R : F \longrightarrow G)$, we pick $[\lambda y \exists x Rxy]$ as our witness. Then by (875.1), it suffices to show:

(A) $\forall x(Fx \rightarrow \exists!y([\lambda y \exists x Rxy]y \,\&\, Rxy))$

(B) $\forall x \forall y(Fx \,\&\, Rxy \rightarrow [\lambda y \exists x Rxy]y)$

(C) $\forall x(Fx \equiv \exists y Rxy)$

Since (C) follows by definition (872.1) from the assumption $HasDomain(R, F)$, it remains only to show (A) and (B).

(A) By GEN, we have to show $Fx \rightarrow \exists!y([\lambda y \exists x Rxy]y \,\&\, Rxy)$. By β-Conversion and a Rule of Substitution, it suffices to show $Fx \rightarrow \exists!y(\exists x Rxy \,\&\, Rxy)$. So assume Fx. Then by (ξ), we know $\exists y Rxy$. So, suppose b is such an object, so that we know Rxb. Now to show $\exists!y(\exists x Rxy \,\&\, Rxy)$, we have to show:

$$\exists y(\exists x Rxy \,\&\, Rxy \,\&\, \forall z(\exists x Rxz \,\&\, Rxz \rightarrow z = y))$$

It suffices to show b is the witness. We already know Rxb, and it follows from this that $\exists x Rxb$. So it remains to show $\forall z(\exists x Rxz \,\&\, Rxz \rightarrow z = y)$. By GEN, it suffices to show $\exists x Rxz \,\&\, Rxz \rightarrow z = b$. So assume $\exists x Rxz$ and Rxz. From the latter and Rxb it follows from (ϑ) and our assumption Fx that $z = b$.

(B) By GEN, we have to show $Fx \,\&\, Rxy \rightarrow [\lambda y \exists x Rxy]y$. So assume Fx and Rxy. But the latter implies $\exists x Rxy$. So by β-Conversion, $[\lambda y\, \exists x Rxy]y$. ⋈

(876.1) By (874.2), we have to show:

(A) $[\lambda xyz\, \mathcal{U}x \,\&\, y =_D z] \dotdiv [\lambda xy\, \mathcal{U}x \,\&\, D!y] \longrightarrow D!$

(B) $HasDomain([\lambda xyz\, \mathcal{U}x \,\&\, y =_D z], [\lambda xy\, \mathcal{U}x \,\&\, D!y])$

(A) By (866.2), we have to show both:

$$[\lambda xyz\,\mathcal{U}x \,\&\, y =_D z]\,|\!:[\lambda xy\,\mathcal{U}x \,\&\, D!y] \longrightarrow D!$$

$$\forall x\forall y\forall z([\lambda x_1 y_1\,\mathcal{U}x_1 \,\&\, D!y_1]xy \,\&\, [\lambda x_1 y_1 z_1\,\mathcal{U}x_1 \,\&\, y_1 =_D z_1]xyz \rightarrow D!z)$$

But the first was established as (863.1). For the second, we have to show, by β-Conversion and a Rule of Substitution:

$$\forall x\forall y\forall z((\mathcal{U}x \,\&\, D!y \,\&\, \mathcal{U}x \,\&\, y =_D z) \rightarrow D!z)$$

By GEN, and simplifying, assume: $\mathcal{U}x$, $D!y$, and $y =_D z$. Then from the latter, by (273.18) that $D!z$.

(B) By (872.3), we have to show:

$$\forall x\forall y([\lambda x_1 y_1\,\mathcal{U}x_1 \,\&\, D!y_1]xy \equiv \exists z([\lambda x_1 y_1 z_1\,\mathcal{U}x_1 \,\&\, y_1 =_D z_1]xyz))$$

By β-Conversion and a Rule of Subsitution, we have to show:

$$\forall x\forall y((\mathcal{U}x \,\&\, D!y) \equiv \exists z(\mathcal{U}x \,\&\, y =_D z))$$

By GEN, we have to show:

$$(\vartheta)\quad (\mathcal{U}x \,\&\, D!y) \equiv \exists z(\mathcal{U}x \,\&\, y =_D z)$$

($\rightarrow$) Assume $\mathcal{U}x \,\&\, D!y$. Then choose our witness to the right-side of (ϑ) to be y. So by $\exists$I, we have to show $\mathcal{U}x \,\&\, y =_D y$. But $\mathcal{U}x$ is true by the first conjunct of our assumption, and $y =_D y$ follows from the second conjunct by (273.30).

($\leftarrow$) Assume $\exists z(\mathcal{U}x \,\&\, y =_D z)$. Suppose that c is such an object, so that we know $\mathcal{U}x \,\&\, y =_D c$. Then it remains only to show $D!y$. But this follows from $y =_D c$, by (273.18). $\bowtie$

(876.2) By (874.1), we have to show:

(A) $\doteq\,\div\,\mathbb{N} \longrightarrow \mathbb{N}$

(B) $HasDomain(\doteq, \mathbb{N})$

(A) By (866.1), we have to show:

$$\doteq\,|\!:\mathbb{N} \longrightarrow \mathbb{N}$$

$$\forall x\forall y(\mathbb{N}x \,\&\, x \doteq y \rightarrow \mathbb{N}y)$$

But the first is just (863.2). For the second, assume $\mathbb{N}x$ and $x \doteq y$. By the latter and theorem (839.2), it follows that $\mathbb{N}y$.

(B) By (872.1), we have to show:

$$\forall x(\mathbb{N}x \equiv \exists y(x \doteq y))$$

and by GEN, that $\mathbb{N}x \equiv \exists y(x \doteq y)$. ($\rightarrow$) Assume $\mathbb{N}x$. Then by (840.3), $x \doteq x$. Hence $\exists y(x \doteq y)$. ($\leftarrow$) Assume $\exists y(x \doteq y)$. Suppose a is such an object, so that we know $x \doteq a$. But by theorem (839.2), $\mathbb{N}x$. ⋈

(876.3) By definition (849), $\mathcal{U}^0$ is $[\lambda\, p_0]$, i.e., by (111.1), p_0, where p_0 is the proposition $\forall x(E!x \rightarrow E!x)$, by a convention in (208). So by (875.3), we have to show:

(A) $p_0 \rightarrow \exists! y([\lambda n\, \mathbb{P}0n]y \,\&\, [\lambda n\, n \doteq 1]y)$

(B) $\forall y(p_0 \,\&\, [\lambda n\, n \doteq 1]y \rightarrow [\lambda n\, \mathbb{P}0n]y)$

(C) $p_0 \equiv \exists y[\lambda n\, n \doteq 1]y$

(A) By the first axiom of propositional logic (38.1), it suffices to show the consequent. By elimination of the restricted variables, β-Conversion, and a Rule of Substitution, we have to show:

$$\exists! y(\mathbb{N}y \,\&\, \mathbb{P}0y \,\&\, \mathbb{N}y \,\&\, y \doteq 1)$$

By simplifying (i.e., removing the otiose third conjunct), the definition of the uniqueness quantifier requires us to show:

$$\exists y(\mathbb{N}y \,\&\, \mathbb{P}0y \,\&\, y \doteq 1 \,\&\, \forall z((\mathbb{N}z \,\&\, \mathbb{P}0z \,\&\, z \doteq 1) \rightarrow z = y)$$

But this is easy to show if we let One be our witness. For we know $\mathbb{N}1$ (827.1), and $\mathbb{P}01$ (826.1). The former implies $1 \doteq 1$, by (840.3). So it remains to show uniqueness. By now familiar reasoning, assume $\mathbb{N}z \,\&\, \mathbb{P}0z \,\&\, z \doteq 1$. Then the third conjunct yields $z = 1$, by (840.1).

(B) Assume p_0 and $[\lambda n\, n \doteq 1]y$, i.e., $\mathbb{N}y \,\&\, y \doteq 1$. But by (840.1), $y \doteq 1$ implies $y = 1$. Now we also know $\mathbb{P}01$, by (826.1). Since $\mathbb{N}1$ (827.1), it follows by β-Conversion that $[\lambda n\, \mathbb{P}0n]1$, i.e., $[\lambda n\, \mathbb{P}0n]y$.

(C) By β-Conversion, a Rule of Substitution, and eliminating the restricted variables, we have to show:

$$p_0 \equiv \exists y(\mathbb{N}y \,\&\, y \doteq 1)$$

($\rightarrow$) By propositional logic, it suffices to show $\exists y(\mathbb{N}y \,\&\, y \doteq 1)$. But this is easy if we let One be our witness, since $\mathbb{N}1$ is a theorem (827.1), and hence, so is $1 \doteq 1$ (840.3). So by &I and $\exists$I, $\exists y(\mathbb{N}y \,\&\, y \doteq 1)$. ($\leftarrow$) p_0 is a theorem. So $\exists y(\mathbb{N}y \,\&\, y \doteq 1) \rightarrow p_0$. ⋈

(877.1) Assume:

(ϑ) $R : F \longrightarrow G$

(ζ) $\forall x(Fx \equiv Hx)$

To show $R : H \longrightarrow G$, it suffices to show, by (875.1):

(A) $\forall x(Hx \rightarrow \exists!y(Gy \,\&\, Rxy))$

(B) $\forall x \forall y(Hx \,\&\, Rxy \rightarrow Gy)$

(C) $HasDomain(R, H)$

(A) By GEN, we need only show the embedded conditional. So assume Hx. Then Fx, by (ζ). Since it follows from (ϑ) by (875.1) *a fortiori* that $R|: F \longrightarrow G$, i.e., that $\forall x(Fx \rightarrow \exists!y(Gy \,\&\, Rxy))$, we can conclude $\exists!y(Gy \,\&\, Rxy)$.

(B) By two applications of GEN, we need only show the embedded conditional. So assume Hx and Rxy. But from Hx, it follows that Fx, by (ζ). Yet (ϑ) implies by (875.1) *a fortiori* that $\forall x \forall y(Fx \,\&\, Rxy \rightarrow Gy)$. Since we have Fx and Rxy, it follows that Gy.

(C) Now (ϑ) implies by (875.1) *a fortiori* that $\forall x(Fx \equiv \exists y Rxy)$. But from this and ($\zeta$) it follows that $\forall x(Hx \equiv \exists y Rxy)$. Hence $HasDomain(R, H)$. ⋈

(877.2) (Exercise)

(877.3) Assume:

(ϑ) $R : p \longrightarrow G$

(ζ) $p \equiv q$

To show $R : p \longrightarrow G$, it suffices to show, by (875.1):

(A) $q \rightarrow \exists!y(Gy \,\&\, Ry)$

(B) $\forall y(q \,\&\, Ry \rightarrow Gy)$

(C) $HasDomain(R, q)$

(A) Assume q. Then p, by (ζ). Since it follows from (ϑ) by (875.1) *a fortiori* that $R|: p \longrightarrow G$, i.e., that $p \rightarrow \exists!y(Gy \,\&\, Ry))$, we can conclude $\exists!y(Gy \,\&\, Ry)$.

(B) By GEN, we show the embedded conditional. So assume q and Ry. But from q, it follows that p, by (ζ). Yet (ϑ) implies by (875.1) *a fortiori* that $\forall y(p \,\&\, Ry \rightarrow Gy)$. Since we have p and Ry, it follows that Gy.

(C) Now (ϑ) implies by (875.1) *a fortiori* that $p \equiv \exists y Ry$. But from this and (ζ) it follows that $q \equiv \exists y Ry$. Hence $HasDomain(R, q)$. ⋈

(878.1) Assume:

(ϑ) $R : F \longrightarrow G$

(ζ) $\forall x(Gx \rightarrow Hx)$

To show $R : F \longrightarrow H$, we have to show, by (875.1):

(A) $\forall x(Fx \rightarrow \exists!y(Hy \,\&\, Rxy))$

(B) $\forall x \forall y(Fx \,\&\, Rxy \rightarrow Hy)$

(C) $\forall x(Fx \equiv \exists y Rxy)$

(A) By GEN, we show the embedded conditional. So assume Fx. Note that it follows from (ϑ) by (875.1) *a fortiori* that $\forall x(Fx \rightarrow \exists!y(Gy \& Rxy))$. This and our assumption Fx imply $\exists!y(Gy \,\&\, Rxy)$, i.e., by the definition of the uniqueness quantifier:

$$\exists y(Gy \,\&\, Rxy \,\&\, \forall z(Gz \,\&\, Rxz \rightarrow z = y))$$

Let a be such an object, so that we know:

(ξ) $Ga \,\&\, Rxa \,\&\, \forall z(Gz \,\&\, Rxz \rightarrow z = a)$

Now we have to show, by definition of the uniqueness quantifier:

$$\exists y(Hy \,\&\, Rxy \,\&\, \forall z(Hz \,\&\, Rxz \rightarrow z = y))$$

But we can show this if we choose a as our witness; by &I and $\exists$I, we have to show: Ha, Rxa, and $\forall z(Hz \,\&\, Rxz \rightarrow z = a)$. From the first conjunct of (ξ) and (ζ), it follows that Ha. Rxa is also known – it's the second conjunct of (ξ). So, finally, assume Hz and Rxz (to show $z = a$). Then since we know Fx by assumption, we have:

$$Fx \,\&\, Rxz$$

But from (ϑ) it follows *a fortiori* by (875.1) that $\forall x \forall y(Fx \,\&\, Rxy \rightarrow Gy)$. Hence from $Fx \,\&\, Rxz$ it follows that Gz. But from this and Rxz, it follows from the third conjunct of (ξ) that $z = a$.

(B) By GEN, we show the embedded conditional. So assume Fx and Rxy. Now we just established that we know $\forall x \forall y(Fx \,\&\, Rxy \rightarrow Gy)$. Hence Gy. So by (ζ), Hy.

(C) This is already known, since it follows from (ϑ) by (875.1) *a fortiori*. ⋈

(878.2) – **(878.3)** (Exercises)

(879.1) We show the first conjunct $\mathbb{P}_{\restriction\mathbb{N}} : \mathbb{N} \longrightarrow \mathbb{N}$, and then the second conjunct *HasRange*$(\mathbb{P}_{\restriction\mathbb{N}}, \mathbb{N}_{+})$.

For the first conjunct, we have to show, by definition (874.1):

(A) $\mathbb{P}_{\restriction\mathbb{N}} \dotdiv \mathbb{N} \longrightarrow \mathbb{N}$

(B) $\forall x(\mathbb{N}x \equiv \exists y \mathbb{P}_{\restriction\mathbb{N}} xy)$

(A) By definition (866.1), we have to show both:

$$\mathbb{P}_{\restriction\mathbb{N}}\,|\!:\mathbb{N}\longrightarrow\mathbb{N}$$

$$\forall x\forall y(\mathbb{N}x \,\&\, \mathbb{P}_{\restriction\mathbb{N}}xy \to \mathbb{N}y)$$

For the first, we can't just cite (864), since that only tells us that $\mathbb{P}\,|\!:\mathbb{N}\longrightarrow\mathbb{N}$. But from this fact and (865.1), it follows that $\mathbb{P}_{\restriction\mathbb{N}}\,|\!:\mathbb{N}\longrightarrow\mathbb{N}$. For the second, assume, by GEN, $\mathbb{N}x$ and $\mathbb{P}_{\restriction\mathbb{N}}xy$. The latter, by definition of $\mathbb{P}_{\restriction\mathbb{N}}$ (828.1) and β-Conversion, implies both $\mathbb{N}x$ and $\mathbb{P}xy$. So by (814.1), $\mathbb{N}y$.

(B) By GEN, we have to show $\mathbb{N}x \equiv \exists y\mathbb{P}_{\restriction\mathbb{N}}xy$:

(→) Assume $\mathbb{N}x$. Then by (818), it follows *a fortiori* that $\exists y\mathbb{P}xy$. Suppose a is such an object, so that we know $\mathbb{P}xa$. Then we know:

$$\mathbb{N}x \,\&\, \mathbb{P}xa$$

Hence, by β-Conversion,

$$[\lambda xy\, \mathbb{N}x \,\&\, \mathbb{P}xy]xa$$

So by definition (828.1), $\mathbb{P}_{\restriction\mathbb{N}}xa$. Hence, $\exists y\mathbb{P}_{\restriction\mathbb{N}}xy$.

(←) Assume $\exists y\mathbb{P}_{\restriction\mathbb{N}}xy$. Suppose a is such an object, so that we know $\mathbb{P}_{\restriction\mathbb{N}}xa$. Then by definition of $\mathbb{P}_{\restriction\mathbb{N}}$ (828.1) and β-Conversion, it follows *a fortiori* that $\mathbb{N}x$.

Finally, we show the second conjunct $HasRange(\mathbb{P}_{\restriction\mathbb{N}}, \mathbb{N}_+)$. By definition (872), we have to show:

(ϑ) $\forall y(\mathbb{N}_+y \equiv \exists x\mathbb{P}_{\restriction\mathbb{N}}xy)$

So by GEN, we have to show $\mathbb{N}_+y \equiv \exists x\mathbb{P}_{\restriction\mathbb{N}}xy$.

(→) Assume $\mathbb{N}_+y$. Then by (843), it follows *a fortiori* that $\exists x(\mathbb{N}x \,\&\, \mathbb{P}xy)$. Let a be such an object, so that we know $\mathbb{N}a \,\&\, \mathbb{P}ay$. It follows by β-Conversion that $[\lambda xy\, \mathbb{N}x \,\&\, \mathbb{P}xy]ay$. So by definition (828.1), $\mathbb{P}_{\restriction\mathbb{N}}ay$, and by $\exists$I, $\exists x\mathbb{P}_{\restriction\mathbb{N}}xy$.

(←) Assume $\exists x\mathbb{P}_{\restriction\mathbb{N}}xy$. Suppose a is such an object, so that we know $\mathbb{P}_{\restriction\mathbb{N}}ay$. By definition (828.1) and β-Conversion, this implies both $\mathbb{N}a$ and $\mathbb{P}ay$. Now to show $\mathbb{N}_+y$, we have to show $y > 0$, i.e., $0 < y$. By (832.1), it suffices to show $\mathbb{P}^*0y$. But this follows from facts we already have, namely, $\mathbb{P}^+0a$ (i.e., $\mathbb{N}a$) and $\mathbb{P}ay$, by (795.3). ⋈

(880.1) We begin by first establishing a preliminary fact. By theorem (850.3), we know that $\forall x\mathcal{U}x$ is a theorem. So $\mathcal{U}y$ is a theorem. Consequently, if Rxy, then $\mathcal{U}y \,\&\, Rxy$, and if $\mathcal{U}y \,\&\, Rxy$, then Rxy. Hence, by definition of $\equiv$ and GEN, the following quantified biconditional is a modally strict theorem:

(ϑ) $\forall x \forall y(Rxy \equiv (\mathcal{U}y \,\&\, Rxy))$

($\rightarrow$) Now suppose $Function^1(R)$. Then by definition (848.1), we know:

(ζ) $\forall x \exists! y Rxy$

Now by (875.1), we have to show:

(A) $\forall x(\mathcal{U}x \rightarrow \exists! y(\mathcal{U}y \,\&\, Rxy))$

(B) $\forall x \forall y(\mathcal{U}x \,\&\, Rxy \rightarrow \mathcal{U}y)$

(C) $\forall x(\mathcal{U}x \equiv \exists y Rxy)$

(A) Note that from from (ζ) and (ϑ), it follows by a Rule of Substitution that $\forall x \exists! y(\mathcal{U}y \& Rxy)$. Moreover, by quantifier law (99.12), we know $\forall \alpha \varphi \rightarrow \forall \alpha(\psi \rightarrow \varphi)$. Hence $\forall x(\mathcal{U}x \rightarrow \exists! y(\mathcal{U}y \,\&\, Rxy))$.

(B) Assume $\mathcal{U}x \,\&\, Rxy$. But Rxy alone implies $\mathcal{U}y$, by (ϑ).

(C) By GEN, we have to show $\mathcal{U}x \equiv \exists y Rxy$. ($\rightarrow$) It follows from ($\zeta$) *a fortiori* that $\exists y Rxy$. Hence $\mathcal{U}x \rightarrow \exists y Rxy$. ($\leftarrow$) We know $\mathcal{U}x$ is a theorem (850.3). Hence $\exists y Rxy \rightarrow \mathcal{U}x$. ⋈

($\leftarrow$) Assume $R : \mathcal{U} \longrightarrow \mathcal{U}$. Then by definition (866.1), it follows *a fortiori* that $R|: \mathcal{U} \longrightarrow \mathcal{U}$, which by (743.1), means:

(ξ) $\forall x(\mathcal{U}x \rightarrow \exists! y(\mathcal{U}y \,\&\, Rxy))$

Now we have to show $\forall x \exists! y Rxy$, so by GEN, we have to show $\exists! y Rxy$. Now by (ξ), we know $\mathcal{U}x \rightarrow \exists! y(\mathcal{U}y \,\&\, Rxy)$, and by theorem (850.3), we know $\mathcal{U}x$. Hence $\exists! y(\mathcal{U}y \,\&\, Rxy)$. But by the modally-strict (ϑ) and a Rule of Substitution, it follows that $\exists! y Rxy$. ⋈

(880.2) (Exercise)

(880.3) (Although this theorem is a special case of (880.2), we give an independent proof.) We begin with a preliminary fact. Since we know $\forall x \mathcal{U}x$ (850.3), it follows that $\mathcal{U}y$ and so we can reason as follows: if Ry, then $\mathcal{U}y \,\&\, Ry$, and if $\mathcal{U}y \,\&\, Ry$, then Ry. Hence, by the definition of $\equiv$ and GEN, we know the following is a modally strict theorem:

(ϑ) $\forall y(Ry \equiv (\mathcal{U}y \,\&\, Ry))$

Now we argue both directions of our theorem.

($\rightarrow$) Assume $Function^0(R)$, i.e., R is a nullary total function. By definition (848.3), this implies $\exists! y Ry$. Now by (875.3), it suffices to show:

(A) $\mathcal{U}^0 \rightarrow \exists! y(\mathcal{U}y \,\&\, Ry)$

(B) $\forall y(\mathcal{U}^0 \,\&\, Ry \rightarrow \mathcal{U}y)$

(C) $\mathcal{U}^0 \equiv \exists y Ry$

(A) We know $\exists! y Ry$, and so by (ϑ) and a Rule of Substitution, $\exists! y(\mathcal{U}y \,\&\, Ry)$. Hence $\mathcal{U}^0 \rightarrow \exists! y(\mathcal{U}y \,\&\, Ry)$.

(B) Assume $\mathcal{U}^0$ and Ry. But Ry alone implies $\mathcal{U}y$, by (ϑ).

(C) We know $\mathcal{U}^0$ is true (850.2) and it follows *a fortiori* from $\exists! y Ry$ that $\exists y Ry$. Hence $\mathcal{U}^0 \equiv \exists y Ry$.

($\leftarrow$) Assume $R : \mathcal{U}^0 \longrightarrow \mathcal{U}$. Then by summary fact (875.3), it follows *a fortiori* that:

$$\mathcal{U}^0 \rightarrow \exists! y(\mathcal{U}y \,\&\, Ry)$$

Now by (848.3), we have to show $\exists! y Ry$. But $\mathcal{U}^0$ is true (850.2). So $\exists! y(\mathcal{U}y \,\&\, Ry)$. But then by the modally strict fact (ϑ) and a Rule of Substitution, $\exists! y Ry$. $\bowtie$

(880.4) Assume $Function^0(R)$. Then by (880.3), we know that $R : \mathcal{U}^0 \longrightarrow \mathcal{U}$. Hence, by definition (874.3), it follows *a fortiori* that $HasDomain(R, \mathcal{U}^0)$. So by definition (872.5):

(ϑ) $\mathcal{U}^0 \equiv \exists y Ry$

Now we have to show $\forall p(p \equiv HasDomain(R, p))$. So by GEN, we show $p \equiv HasDomain(R, p)$. ($\rightarrow$) Assume p. Then since both p is true (by hypothesis) and $\mathcal{U}^0$ is true (850.2), $p \equiv \mathcal{U}^0$. But from this and (ϑ), it follows that $p \equiv \exists y Ry$. Hence by definition (872.5), $HasDomain(R, p)$. ($\leftarrow$) Assume $HasDomain(R, p)$. Then by definition (872.5), $p \equiv \exists y Ry$. But from this and (ϑ) it follows that $\mathcal{U}^0 \equiv p$. Since $\mathcal{U}^0$ is true (850.2), it follows that p. $\bowtie$

(882.1) Assume $R : F \underset{\text{onto}}{\longrightarrow} G$. Then by (881.4), we know both:

(ϑ) $R : F \longrightarrow G$

(ξ) $\forall y(Gy \rightarrow \exists x Rxy)$

To show $HasRange(R, G)$, we have to show, by (872.2) $\forall y(Gy \equiv \exists x Rxy)$. Given ($\xi$), it remains only to show $\forall y(\exists x Rxy \rightarrow Gy)$. So, by GEN, assume $\exists x Rxy$. Let b be such an object, so that we know Rby. Then by $\exists$I, $\exists y Rby$. But from (ϑ) and its definition (874.1), we know that $HasDomain(R, F)$, i.e., by (872.1), $\forall x(Fx \equiv \exists y Rxy)$. Hence Fb. But from (ϑ) and the summary fact (875.1), we know *a fortiori* that $\forall x(Fx \,\&\, Rxy \rightarrow Gy)$. Since we've established Fb and know Rby by assumption, it follows that Gy. $\bowtie$

(882.2) (Exercise)

(882.3) (Although this theorem is a special case of (882.2), we give an independent proof.) Assume $R : p \underset{\text{onto}}{\longrightarrow} G$. Then by definition (881.6), we know:

(ϑ) $R : p \longrightarrow G$

(ξ) $\forall y(Gy \rightarrow Ry)$

To show $HasRange(R, G)$, we have to show, by (872.6), $\forall y(Gy \equiv Ry)$. Given (ξ), it remains to show $\forall y(Ry \rightarrow Gy)$. So assume Ry. Then $\exists yRy$. But from (ϑ) and its definition (874.3), we know that $HasDomain(R, p)$, i.e., by (872.5), $p \equiv \exists yRy$. Hence p. But from (ϑ) and the summary fact (875.1), we know *a fortiori* that $\forall y(p \,\&\, Ry \rightarrow Gy)$. Since we established p and know Ry by assumption, Gy. ⋈

(885.1) ($\rightarrow$) Assume $\upharpoonright$-$function^1(R)$. So by definition (884.1), $\exists F \exists G(R : F \longrightarrow G \,\&\, \neg Function^1(R))$. Let P and Q be such properties, so that we know both:

(ϑ) $R : P \longrightarrow Q$

(ζ) $\neg Function^1(R)$

Given (ϑ), it remains only show $\neg\forall x(Px \equiv \mathcal{U}x)$, by &I and $\exists$I. For reductio, assume $\forall x(Px \equiv \mathcal{U}x)$. From this assumption, (ϑ), and (877.1), it follows that $R : \mathcal{U} \longrightarrow Q$. Moreover, since we know $\forall x \mathcal{U}x$, we also know that $\forall x(Qx \rightarrow \mathcal{U}x)$. Hence by (878.1), $R : \mathcal{U} \longrightarrow \mathcal{U}$. So by (880.1), $Function^1(R)$, which contradicts (ζ).

($\leftarrow$) Assume $\exists F \exists G(R : F \longrightarrow G \,\&\, \neg\forall x(Fx \equiv \mathcal{U}x))$. Suppose P and Q are such properties, so that we know both:

(ξ) $R : P \longrightarrow Q$

(ω) $\neg\forall x(Px \equiv \mathcal{U}x)$

To show $\upharpoonright$-$function^1(R)$, we have to show, by (884):

$$\exists F \exists G(R : F \longrightarrow G \,\&\, \neg Function^1(R))$$

If we pick our witnesses to be P and Q, it remains only to show $\neg Function^1(R)$. Assume, for reductio, $Function^1(R)$. Then by (880.1), $R : \mathcal{U} \longrightarrow \mathcal{U}$. This implies, by (874.1), $HasDomain(R, \mathcal{U})$, i.e., by (872.1):

(A) $\forall x(\mathcal{U}x \equiv \exists yRxy)$

But from (ξ), it also follows by (874.1) that $HasDomain(R, P)$, i.e., by (872.1):

(B) $\forall x(Px \equiv \exists yRxy)$

From (A) and (B) it follows that $\forall x(Px \equiv \mathcal{U}x)$, which contradicts ($\omega$). ⋈

(885.2) (Exercise)

(885.3) (Although this theorem is a special case of theorem (885.2), we give an independent proof.) ($\rightarrow$) Suppose $\upharpoonright$-$function^0(R)$. Then by definition (884.3):

$$\exists p \exists G(R : p \longrightarrow G \,\&\, \neg Function^0(R))$$

Suppose q_1 and P are such a proposition and property, so that we know:

(ϑ) $R : q_1 \longrightarrow P \,\&\, \neg Function^0(R)$

Given the first conjunct of (ϑ), it remains only to show $q_1 \not\equiv \mathcal{U}^0$, by &I and $\exists$I. For reductio, suppose $q_1 \equiv \mathcal{U}^0$. Then from this assumption, the first conjunct of (ϑ) and (877.3), it follows that $R : \mathcal{U}^0 \longrightarrow P$. Now since we know $\forall x \mathcal{U} x$, we also know $\forall x(Px \rightarrow \mathcal{U} x)$. Hence by (878.3), $R : \mathcal{U}^0 \longrightarrow \mathcal{U}$. So by (880.3), $Function^0(R)$, which contradicts the second conjunct of (ϑ).

($\leftarrow$) Suppose $\exists p \exists G(R : p \longrightarrow G \,\&\, p \not\equiv \mathcal{U}^0)$. Let q_1 and P be such a proposition and property, so that we know:

(ξ) $R : q_1 \longrightarrow P \,\&\, q_1 \not\equiv \mathcal{U}^0$

By the definition of a restricted function (884.3), it remains only to show that R is not a total function, i.e., by (848.3), $\neg\exists! y Ry$. From the second conjunct of (ξ) and the truth of $\mathcal{U}^0$ (850.2), we know $\neg q_1$. Separately, from the first conjunct of (ξ), it follows from (874.3) that $HasDomain(R, q_1)$. Hence by definition (872.5), it follows that $\neg\exists y Ry$. But if nothing exemplifies R, nothing uniquely exemplifies R, i.e., $\neg\exists! y Ry$. ⋈

(885.4) ($\rightarrow$) Assume $\restriction$-$function^0(R)$. Then by (885.3), $\exists p \exists G(R : p \longrightarrow G \,\&\, p \not\equiv \mathcal{U}^0)$. Let q_1 and P be such a proposition and property, so that we know $R : q_1 \longrightarrow P$ and $q_1 \not\equiv \mathcal{U}^0$. Since q_1 is not equivalent to the known truth $\mathcal{U}^0$ (850.2), it follows that $\neg q_1$. Since $R : q_1 \longrightarrow P$ implies R has q_1 as a domain (874.3), it follows by definition of $HasDomain(R, q_1)$ (872.5) that $q_1 \equiv \exists y Ry$. Hence $\neg\exists y Ry$.

($\leftarrow$) Assume $\neg\exists y Ry$. Let p be any false proposition, say $\overline{p_0}$; and let G be any property whatsoever. Then if we can show both:

(A) $R : p \longrightarrow G$

(B) $p \not\equiv \mathcal{U}^0$

then by &I and $\exists$I, it follows that $\exists p \exists G(R : p \longrightarrow G \,\&\, p \not\equiv \mathcal{U}^0)$, which by (885.3) establishes that $\restriction$-$function^0(R)$.

(A) By (875.3), we have to show:

(i) $p \rightarrow \exists y(Gy \,\&\, Ry)$

(ii) $\forall y(p \,\&\, Ry \rightarrow Gy)$

(iii) $p \equiv \exists y Ry$

(i) Since we know $\neg p$ by hypothesis, it follows by failure of the antecedent that $p \to \exists y(Gy \,\&\, Ry)$.

(ii) Since we know $\neg p$ by hypothesis, it follows that $p \,\&\, Ry \to Gy$, again by failure of the antecedent. So by GEN, $\forall y(p \,\&\, Ry \to Gy)$.

(iii) We know both $\neg p$ and $\neg\exists y Ry$ by hypothesis. So p and $\exists y Ry$ have the same truth value, i.e., $p \equiv \exists y Ry$.

(B) We know both $\neg p$ by hypothesis and $\mathcal{U}^0$ by theorem (850.2). Hence p and $\mathcal{U}^0$ have different truth values, i.e., $p \not\equiv \mathcal{U}^0$. ⋈

(886.1) Let R be the relation:

$$[\lambda xy\, D!x \,\&\, y \doteq 0]$$

To show $\restriction$-*function*(R), it suffices to show, by (885.2), that:

$$\exists F \exists G(R : F \longrightarrow G \,\&\, \neg\forall x(Fx \equiv \mathcal{U}x))$$

If we choose F to be $D!$ and G to be $\mathbb{N}$, then by &I and ∃I, it suffices to show:

(A) $R : D! \longrightarrow \mathbb{N}$

(B) $\neg\forall x(D!x \equiv \mathcal{U}x)$

(A) By (875.1), we have to show:

(i) $\forall x(D!x \to \exists! y(\mathbb{N}y \,\&\, Rxy))$

(ii) $\forall x \forall y(D!x \,\&\, Rxy \to \mathbb{N}y)$

(iii) $\forall x(D!x \equiv \exists y Rxy)$

(i) By GEN, we show the embedded conditional. So assume $D!x$. Then we pick Zero to be our witness. By the definition of the uniqueness quantifier and ∃I, we have to show:

$$\mathbb{N}0 \,\&\, Rx0 \,\&\, \forall z(\mathbb{N}z \,\&\, Rxz \to z = 0)$$

But $\mathbb{N}0$ is a theorem, and implies $0 \doteq 0$, by (840.3). Hence, conjoining some facts we have established, we obtain:

$$D!x \,\&\, 0 \doteq 0$$

But this is all we need to show $Rx0$, by β-Conversion and definition of R. It remains therefore to show $\forall z(\mathbb{N}z \& Rxz \to z = 0)$. By GEN, we show the embedded conditional. So assume $\mathbb{N}z \,\&\, Rxz$. The second conjunct implies, by definition of R and β-Conversion, that $D!x \,\&\, z \doteq 0$. But then $z \doteq 0$ implies $z = 0$, by (840.1).

(ii) By two applications of GEN, we need only show the embedded conditional. So assume $D!x$ and Rxy. Then by the latter, it follows from the definition of R and β-Conversion that:

$D!x \,\&\, y \doteq 0$

The second conjunct implies $y = 0$, by now familiar reasoning. So from the fact that $\mathbb{N}0$, it follows that $\mathbb{N}y$.

(iii) By GEN, we have to show $D!x \equiv \exists y Rxy$:

($\rightarrow$) Assume $D!x$. Independently, we know $\mathbb{N}0$ and consequently that $0 \doteq 0$. So, we have established $D!x \,\&\, 0 \doteq 0$. So by β-Conversion and definition of R, $Rx0$. Hence, $\exists y Rxy$.

($\leftarrow$) Assume $\exists y Rxy$. Let a be the witness, so that we know Rxa. Then by definition of R and β-Conversion, it follows *a fortiori* that $D!x$.

(B) Since there are abstract objects, let b be one, so that we know $A!b$. Hence, by (222.3) $\neg D!b$. But we know $\mathcal{U}b$, by (850.3). So $\neg(D!b \equiv \mathcal{U}b)$. Hence, $\exists x \neg(D!x \equiv \mathcal{U}x)$, i.e., $\neg\forall x(D!x \equiv \mathcal{U}x)$. ⋈

(886.2) We prove this for the case $n \geq 1$, and prove the case $n = 0$ in (886.3). Let R be the relation:

$$[\lambda x_1 \ldots x_n y\; D!x_1 \,\&\, \ldots \,\&\, D!x_n \,\&\, y \doteq 0]$$

To show $\upharpoonright$-*function*(R), it suffices to show, by (885.2), that:

$$\exists S^n \exists G(R : S^n \longrightarrow G \,\&\, \neg\forall x_1 \ldots \forall x_n(S^n x_1 \ldots x_n \equiv \mathcal{U}^n x_1 \ldots x_n))$$

If we choose S^n to be $D!^{\times n}$ and G to be $\mathbb{N}$, then by &I and $\exists$I, it suffices to show:

(A) $R : D!^{\times n} \longrightarrow \mathbb{N}$

(B) $\neg\forall x_1 \ldots \forall x_n(D!^{\times n} x_1 \ldots x_n \equiv \mathcal{U}^n x_1 \ldots x_n)$

(A) By (875.2), we have to show:

(i) $\forall x_1 \ldots \forall x_n(D!^{\times n} x_1 \ldots x_n \rightarrow \exists! y(\mathbb{N}y \,\&\, Rx_1 \ldots x_n y))$

(ii) $\forall x_1 \ldots \forall x_n \forall y(D!^{\times n} x_1 \ldots x_n \,\&\, Rx_1 \ldots x_n y \rightarrow \mathbb{N}y)$

(iii) $\forall x_1 \ldots \forall x_n(D!^{\times n} x_1 \ldots x_n \equiv \exists y Rx_1 \ldots x_n y)$

(i) By n applications of GEN, we show only the embedded conditional. Assume $D!^{\times n} x_1 \ldots x_n$. Then we pick Zero to be our witness. By the definition of the uniqueness quantifier and $\exists$I, we have to show:

$$\mathbb{N}0 \,\&\, Rx_1 \ldots x_n 0 \,\&\, \forall z(\mathbb{N}z \,\&\, Rx_1 \ldots x_n z \rightarrow z = 0)$$

But $\mathbb{N}0$ is a theorem, and implies $0 \doteq 0$, by (840.3). Now our assumption $D!^{\times n} x_1 \ldots x_n$ implies, by definition (883.2) and β-Conversion:

$$D!x_1 \,\&\, \ldots \,\&\, D!x_n$$

Hence, conjoining some facts we have established, we obtain:

$$D!x_1 \,\&\, \ldots \,\&\, D!x_n \,\&\, 0 \doteq 0$$

But this is all we need to show $Rx_1 \ldots x_n 0$, by β-Conversion and definition of R. It remains therefore to show $\forall z(\mathbb{N}z \,\&\, Rx_1 \ldots x_n z \rightarrow z = 0)$ and, by GEN, the embedded conditional. So assume $\mathbb{N}z \,\&\, Rx_1 \ldots x_n z$. The second conjunct implies, by definition of R and β-Conversion, that $D!x_1 \,\&\, \ldots \,\&\, D!x_n \,\&\, z \doteq 0$. But then the last conjunct implies $z = 0$, by (840.1).

(ii) By GEN, we show the embedded conditional. Assume $D!^{\times n}x_1 \ldots x_n$ and $Rx_1 \ldots x_n y$. Then by the latter, it follows from the definition of R and β-Conversion that:

$$D!x_1 \,\&\, \ldots \,\&\, D!x_n \,\&\, y \doteq 0$$

The last conjunct implies $y = 0$, by now familiar reasoning. So from the fact that $\mathbb{N}0$, it follows that $\mathbb{N}y$.

(iii) By GEN, we have to show:

$$D!^{\times n}x_1 \ldots x_n \equiv \exists y Rx_1 \ldots x_n y$$

($\rightarrow$) Assume $D!^{\times n}x_1 \ldots x_n$. Then it follows by definition (883.2) and β-Conversion that $D!x_1 \,\&\, \ldots \,\&\, D!x_n$. Independently, we know $\mathbb{N}0$ and consequently that $0 \doteq 0$. So, conjoining some things we know:

$$D!x_1 \,\&\, \ldots \,\&\, D!x_n \,\&\, 0 \doteq 0$$

So by β-Conversion and definition of R, $Rx_1 \ldots x_n 0$. Hence, $\exists y Rx_1 \ldots x_n y$.

($\leftarrow$) Assume $\exists y Rx_1 \ldots x_n y$. Let a be the witness, so that we know $Rx_1 \ldots x_n a$. Then by definition of R and β-Conversion, it follows *a fortiori* that $D!x_1 \,\&\, \ldots \,\&\, D!x_n$. So by definition (883.2) and β-Conversion, $D!^{\times n}!x_1 \ldots x_n$.

(B) Since there are abstract objects, let a be one. Then we know $A!a \,\&\, \ldots \,\&\, A!a$. Hence, by (222.3) $\neg D!a \,\&\, \ldots \,\&\, \neg D!a$. *A fortiori*, $\neg(D!a \,\&\, \ldots \,\&\, D!a)$, i.e., by β-Conversion and definition (883.2), $\neg(D!^{\times n}a \ldots a)$. But we know $\mathcal{U}^n a \ldots a$, by (850.3) So:

$$\neg(D!^{\times n}a \ldots a \equiv \mathcal{U}^n a \ldots a)$$

Hence,

$$\exists x_1 \ldots \exists x_n \neg(D!^{\times n}x_1 \ldots x_n \equiv \mathcal{U}^n x_1 \ldots x_n)$$

I.e.,

$$\neg\forall x_1 \ldots \forall x_n(D!^{\times n}x_1 \ldots x_n \equiv \mathcal{U}^n x_1 \ldots x_n) \qquad \bowtie$$

(886.3) Let R be any unexemplified property. Then by (885.4), $\upharpoonright$-*function*$^0(R)$. $\bowtie$

(887.1) – **(887.4)** (Exercises)

(888.1) Assume $\upharpoonright$-*function*$^0(R)$. Then by (885.3), $\exists p \exists G(R : p \longrightarrow G \,\&\, p \not\equiv \mathcal{U}^0)$. Let q_1 and P be such a proposition and property, so that we know:

(ϑ) $R : q_1 \longrightarrow P \,\&\, q_1 \not\equiv \mathcal{U}^0$

By GEN, we have to show $HasDomain(R, p) \equiv \neg p$. $(\rightarrow)$ Assume $HasDomain(R, p)$. Then by definition (872.5), $p \equiv \exists y Ry$. But by (885.4), $\neg\exists y Ry$. Hence $\neg p$. $(\leftarrow)$ Assume $\neg p$. But we know $\neg q_1$, since the second conjunct of (ϑ) asserts that q_1 fails to be equivalent to the truth $\mathcal{U}^0$. So $p \equiv q_1$. From this fact and the first conjunct of (ϑ), it follows by (877.3) that $R : p \longrightarrow P$. Hence, by definition (874.3), $HasDomain(R, p)$. $\bowtie$

(888.2) Assume $\upharpoonright$-*function*$^0(R)$. Then by (885.3), $\exists p \exists G(R : p \longrightarrow G \,\&\, p \not\equiv \mathcal{U}^0)$. Let q_1 and P be such a proposition and property, so that we know:

(ϑ) $R : q_1 \longrightarrow P \,\&\, q_1 \not\equiv \mathcal{U}^0$

By GEN, we have to show $HasRange(R, H) \equiv \neg\exists y Hy$.

$(\rightarrow)$ Assume $HasRange(R, H)$, i.e., by definition (872.6), $\forall y(Hy \equiv Ry)$. But the fact that R is a nullary restricted function implies $\neg\exists y Ry$, by theorem (885.4). Hence $\neg\exists y Hy$.

$(\leftarrow)$ Assume $\neg\exists y Hy$. To show $HasRange(R, H)$, we have to show, by GEN, $Hy \equiv Ry$:

- $(\rightarrow)$ By our assumption $\neg\exists y Hy$, we know $\forall y \neg Hy$. Hence $\neg Hy$. So $Hy \rightarrow Ry$, by failure of the antecedent.
- $(\leftarrow)$ Since R is a nullary restricted function, we know $\neg\exists y Ry$, by theorem (885.4). So $\forall y \neg Ry$ and, hence, $\neg Ry$. So by failure of the antecedent, $Ry \rightarrow Hy$. $\bowtie$

(890) (Exercise)

(892) Assume $\overline{f}$ is any n-ary function such that $\exists y \overline{f} x_1 \ldots x_n y$. Now by (890), we know that $\overline{f}$ is either total or restricted. If $\overline{f}$ is total, then by (848.2), it follows immediately that $\exists! y \overline{f} x_1 \ldots x_n y$. If $\overline{f}$ is a nullary restricted function, then the result follows immediately by failure of the antecedent. If $\overline{f}$ is a restricted function with $n \geq 1$, then by (889), $\exists S^n \exists G(\overline{f} : S^n \longrightarrow G)$. Suppose A^n and Q are such a relation and property, so that we know $\overline{f} : A^n \longrightarrow Q$. Hence, by (875.2), it follows *a fortiori* that:

(ϑ_1) $\forall x_1 \ldots \forall x_n(A^n x_1 \ldots x_n \rightarrow \exists! y(Qy \,\&\, \overline{f} x_1 \ldots x_n y))$

(ϑ_2) $\forall x_1 \ldots \forall x_n \forall y(A^n x_1 \ldots x_n \mathbin{\&} \overline{f} x_1 \ldots x_n y \rightarrow Qy)$

(ϑ_3) $\forall x_1 \ldots \forall x_n(A^n x_1 \ldots x_n \equiv \exists y \overline{f} x_1 \ldots x_n y)$

Our assumption $\exists y \overline{f} x_1 \ldots x_n y$ implies $A^n x_1 \ldots x_n$ by (ϑ_3). From this and (ϑ_1), it follows that $\exists! y(Qy \mathbin{\&} \overline{f} x_1 \ldots x_n y))$, i.e.,

$$\exists y(Qy \mathbin{\&} \overline{f} x_1 \ldots x_n y \mathbin{\&} \forall z(Qz \mathbin{\&} \overline{f} x_1 \ldots x_n z \rightarrow z = y))$$

Suppose a is such an object, so that we know:

(ζ) $Qa \mathbin{\&} \overline{f} x_1 \ldots x_n a \mathbin{\&} \forall z(Qz \mathbin{\&} \overline{f} x_1 \ldots x_n z \rightarrow z = a)$

Since the second conjunct is $\overline{f} x_1 \ldots x_n a$, it remains only to show $\forall z(\overline{f} x_1 \ldots x_n z \rightarrow z = a)$, since from the conjunction of these two, it follows by $\exists$I and the definition of the uniqueness quantifier that $\exists! y \overline{f} x_1 \ldots x_n y$. So assume $\overline{f} x_1 \ldots x_n z$. Then from the previously established fact that $A^n x_1 \ldots x_n$, it follows from (ϑ_2) that Qz. But then Qz and $\overline{f} x_1 \ldots x_n z$ imply, by the third conjunct of (ζ) that $z = a$. ⋈

(894.2) (Exercise)

(895) By GEN, it suffices to show $Op^n(R) \rightarrow \Box Op^n(R)$. So assume $Op^n(R)$. Then, by (894.1), $Rigid(R)$ and $R : \mathbb{N}^{\times n} \longrightarrow \mathbb{N}$. By now familiar modal principles, it suffices to show both (a) $\Box Rigid(R)$ and (b) $\Box(R : \mathbb{N}^{\times n} \longrightarrow \mathbb{N})$.

(a) The definition of $Rigid(R)$ in (571.1) tells us that our assumption $Rigid(R)$ implies both $R{\downarrow}$ and $\Box \forall x_1 \ldots \forall x_n(R^n x_1 \ldots x_n \rightarrow \Box R^n x_1 \ldots x_n)$. But the former is necessary, since it is the modal closure of an instance of axiom (39.2), and the latter is necessary by the 4 schema. Hence $\Box R{\downarrow}$ and $\Box\Box \forall x_1 \ldots \forall x_n(R^n x_1 \ldots x_n \rightarrow \Box R^n x_1 \ldots x_n)$, it follows by standard modal reasoning and definition (571.1) that $\Box Rigid(R)$.

(b) From our assumption that $R : \mathbb{N}^{\times n} \longrightarrow \mathbb{N}$, it follows by definition (862.1):

(ϑ) $\forall x1 \ldots \forall x_n(\mathbb{N}^{\times n} x_1 \ldots x_n \rightarrow \exists! y(\mathbb{N} y \mathbin{\&} R x_1 \ldots x_n y))$

But suppose this isn't necessary, for reductio. Then:

$$\neg \Box \forall x_1 \ldots \forall x_n(\mathbb{N}^{\times n} x_1 \ldots x_n \rightarrow \exists! y(\mathbb{N} y \mathbin{\&} R x_1 \ldots x_n y))$$

i.e.,

$$\Diamond \exists x_1 \ldots \exists x_n(\mathbb{N}^{\times n} x_1 \ldots x_n \mathbin{\&} \neg \exists! y(\mathbb{N} y \mathbin{\&} R x_1 \ldots x_n y))$$

So by n applications of CBF$\Diamond$:

$$\exists x_1 \ldots \exists x_n \Diamond(\mathbb{N}^{\times n} x_1 \ldots x_n \mathbin{\&} \neg \exists! y(\mathbb{N} y \mathbin{\&} R x_1 \ldots x_n y))$$

from which it follows that:

$$\exists x_1 \ldots \exists x_n(\Diamond \mathbb{N}^{\times n} x_1 \ldots x_n \mathbin{\&} \Diamond \neg \exists! y(\mathbb{N} y \mathbin{\&} R x_1 \ldots x_n y)))$$

Let $a_1, \ldots, a_n$ be arbitrary such objects, so that we know:

(ζ) $\Diamond\mathbb{N}^{\times n} a_1 \ldots a_n \mathbin{\&} \Diamond\neg\exists!y(\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y)$

Our strategy now is to show that $\Box\exists!y(\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y)$which contradicts the 2nd conjunct of (ζ). Now the first conjunct of (ζ) implies $\mathbb{N}^{\times n} a_1 \ldots a_n$, by n applications of the rigidity of $\mathbb{N}$ (809.2) and (172.1). (As an exercise, prove this by induction.) But from this and (ϑ):

(ξ) $\exists!y(\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y)$

Note that the following is a modal fact, for arbitrary φ (174.6):

$$\Box\forall\alpha(\varphi \to \Box\varphi) \to (\exists!\alpha\varphi \to \Box\exists!\alpha\varphi)$$

So let α be y and let φ be $\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y$. So we know, as an instance of our modal fact that:

$$\begin{aligned}&\Box\forall y((\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y) \to \Box(\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y)) \to \\ &\quad (\exists!y(\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y) \to \Box\exists!y(\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y))\end{aligned}$$

Since we already have the antecedent of the consequent as (ξ), and our goal is to show the consequent of the consequent, it remains only to show the antecedent, i.e.,

$$\Box\forall y((\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y) \to \Box(\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y))$$

But by RN and GEN, we need only show, by a modally strict proof, that:

$$(\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y) \to \Box(\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y)$$

So assume $\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y$. Then by the rigidity of $\mathbb{N}$, it follows from the 1st conjunct that $\Box\mathbb{N}y$. And by the rigidity of R, which we established in (a) above, it follows from the 2nd conjunct that $\Box Ra_1 \ldots a_n y$. Hence $\Box(\mathbb{N}y \mathbin{\&} Ra_1 \ldots a_n y)$. Contradiction. ⋈

(896) (Exercise)

(897.2) Fix n and m. By definition, we have to show:

(A) $Rigid(\mathcal{C}^{n'}_m)$

(B) $\mathcal{C}^{n'}_m : \mathbb{N}^{\times n} \longrightarrow \mathbb{N}$

(A) Clearly $\mathcal{C}^{n'}_m$ exists by our theory of definitions and the fact that its definiens is a core λ-expression (9.2), i.e., the λ doesn't bind any variable that occurs in encoding position (9.1) in the matrix. So it remains so show:

$$\Box\forall x_1 \ldots x_n \forall y(\mathcal{C}^{n'}_m x_1 \ldots x_n y \to \Box\mathcal{C}^{n'}_m x_1 \ldots x_n y)$$

By GEN and RN, it suffices to show $\mathcal{C}^{n'}_m x_1 \ldots x_n y \to \Box\mathcal{C}^{n'}_m x_1 \ldots x_n y$. So assume $\mathcal{C}^{n'} x_1 \ldots x_n y$. Then by definition of $\mathcal{C}^{n'}_m$ and λ-Conversion, $\mathbb{N}x_1 \,\&\, \ldots \mathbb{N}x_n \,\&\, y \doteq m$. But $\mathbb{N}x_i$ implies $\Box\mathbb{N}x_i$ $(1 \leq i \leq n)$, by (809.1). And we already know that $\doteq$ is rigid (840.6). So $\Box y \doteq m$. Conjoining our results, we have $\Box(\mathbb{N}x_1 \& \ldots \& \mathbb{N}x_n \& y \doteq m)$. So by λ-Conversion:

$$\Box[\lambda x_1 \ldots x_n y\; \mathbb{N}x_1 \,\&\, \ldots \mathbb{N}x_n \,\&\, y \doteq n]x_1 \ldots x_n y$$

i.e., $\Box\mathcal{C}^{n'}_m x_1 \ldots x_n y$, by Rule $=_{df}$I.

(B) (exercise). ⋈

(898.4) (Exercise)

(899.1) We prove this by cases. *Case 1* $(n \geq 1)$. Assume $Op^n(R)$, so that by (894.1) we know $R : \mathbb{N}^{\times n} \longrightarrow \mathbb{N}$. Hence, by (875.2), we know *a fortiori*:

(ϑ) $\forall x_1 \ldots \forall x_n \forall y(\mathbb{N}^{\times n} x_1 \ldots x_n \,\&\, Rx_1 \ldots x_n y \to \mathbb{N}y)$

(ξ) $\forall x_1 \ldots \forall x_n(\mathbb{N}^{\times n} x_1 \ldots x_n \equiv \exists y Rx_1 \ldots x_n y)$

Now assume $Rx_1 \ldots x_n y$. Hence $\exists y Rx_1 \ldots x_n y$. Then by ($\xi$), it follows that:

(ζ) $\mathbb{N}^{\times n} x_1 \ldots x_n$

So by definition of the Cartesian product (883.2) and β-Conversion, it follows that $\mathbb{N}x_1 \,\&\, \ldots \,\&\, \mathbb{N}x_n$. Moreover, it follows from (ζ) and our assumption $Rx_1 \ldots x_n y$ that $\mathbb{N}y$, by (ϑ).

Case 2 $(n = 0)$. We want to show:

$$Op^0(R) \to \forall y(Ry \to \mathbb{N}y)$$

So, assume $Op^0(R)$. By (894.1), $Rigid(R) \,\&\, R : \mathbb{N}^{\times 0} \longrightarrow \mathbb{N}$. Then by (882.3), $HasRange(R, \mathbb{N})$. And by (872.6), this implies that $\forall y(\mathbb{N}y \equiv Ry)$. *A fortiori*, $\forall y(Ry \to \mathbb{N}y)$. ⋈

(899.2) Assume $Op^n(R)$, so that by (894.1) we know both that R is rigid and that $R : \mathbb{N}^{\times n} \longrightarrow \mathbb{N}$. If $n \geq 1$, then $\exists S^n \exists G(R : S^n \longrightarrow G)$; if $n = 0$, $\exists p \exists G(R : p \longrightarrow G)$. So $function^n(R)$, by (889). We use this fact to prove the consequent of our theorem. By GEN, we prove both directions of the biconditional $Rx_1 \ldots x_n y \equiv y = \imath y Rx_1 \ldots x_n y)$. ($\to$) Assume $Rx_1 \ldots x_n y$. Since R is an n-ary function, it follows by by (892) that y is unique. Hence, we know:

(ϑ) $Rx_1 \ldots x_n y \,\&\, \forall z(Rx_1 \ldots x_n z \to z = y)$

Since R is rigid, it follows from definition (571.1) that:

$$R{\downarrow} \,\&\, \Box\forall x_1 \ldots \forall x_n \forall z(Rx_1 \ldots x_n z \to \Box Rx_1 \ldots x_n z)$$

By $n+1$ applications of CBF and n applications of a Rule of Substitution, the 2nd conjunct of this claim implies:

$$(\zeta)\ \ \forall x_1 \ldots \forall x_n \forall z \Box(Rx_1 \ldots x_n z \to \Box Rx_1 \ldots x_n z)$$

If we instantiate this to $x_1, \ldots, x_n$ and y, we have:

$$(\xi)\ \ \Box(Rx_1 \ldots x_n y \to \Box Rx_1 \ldots x_n y)$$

If we apply the T schema to (ξ), then the result and our assumption yield $\Box Rx_1 \ldots x_n y$. Hence $\mathcal{A}Rx_1 \ldots x_n y$. Put this aside for the moment and now instantiate just the first n quantifiers of (ζ) to $x_1, \ldots, x_n$, so that we obtain:

$$\forall z \Box(Rx_1 \ldots x_n z \to \Box Rx_1 \ldots x_n z)$$

But if we let φ in theorem (174.2) be $Rx_1 \ldots x_n z$ and apply GEN to the resulting instance, we know:

$$\forall z[\Box(Rx_1 \ldots x_n z \to \Box Rx_1 \ldots x_n z) \to (\mathcal{A}Rx_1 \ldots x_n z \equiv Rx_1 \ldots x_n z)]$$

Hence it follows from our last two results by an axiom of predicate logic (39.3) that:

$$\forall z(\mathcal{A}Rx_1 \ldots x_n z \equiv Rx_1 \ldots x_n z)$$

From this and the second conjunct of (ϑ) it follows by predicate logic that $\forall z(\mathcal{A}Rx_1 \ldots x_n z \to z = y)$. So, we may assemble what we have established as:

$$\mathcal{A}Rx_1 \ldots x_n y \ \&\ \forall z(\mathcal{A}Rx_1 \ldots x_n z \to z = y)$$

Hence by an alphabetic variant of of the modally strict version of Hintikka's schema (148), it follows that:

$$y = \imath y Rx_1 \ldots x_n y$$

($\leftarrow$) (Exercise) $\bowtie$

(899.3) We prove this by cases. *Case 1* ($n \geq 1$). Assume $Op^n(R)$. By eliminating our restricted variables, we have to show:

$$\forall x_1 \ldots \forall x_n((\mathbb{N}x_1 \ \&\ \ldots\ \&\ \mathbb{N}x_n) \to \exists! y(\mathbb{N}y \ \&\ Rx_1 \ldots x_n y))$$

So assume $\mathbb{N}x_1 \ \&\ \ldots\ \&\ \mathbb{N}x_n$. Since $Op^n(R)$, we know by (875.1) *a fortiori* that:

$$\forall x_1 \ldots \forall x_n(\mathbb{N}^{\times n} x_1 \ldots x_n \to \exists! y(\mathbb{N}y \ \&\ Rx_1 \ldots x_n y))$$

Instantiating to $x_1, \ldots, x_n$, it follows that:

$$\mathbb{N}^{\times n} x_1 \ldots x_n \to \exists! y(\mathbb{N}y \ \&\ Rx_1 \ldots x_n y)$$

But our assumption $\mathbb{N}x_1$ & … & $\mathbb{N}x_n$ implies the antecedent, by (883.2) and β-Conversion. Hence $\exists!y(\mathbb{N}y \,\&\, Rx_1 \ldots x_n y)$.

Case 2 ($n = 0$). We have to show:

$$Op^0(R) \to \exists!kRk$$

So assume $Op^0(R)$. Then by (894.1), $R : \mathbb{N}^{\times 0} \longrightarrow \mathbb{N}$. By (883.3), $\mathbb{N}^{\times 0} = \mathcal{U}^0$. By (875.3), $\mathcal{U}^0 \to \exists!y(\mathbb{N}y \,\&\, Ry)$. But since $\mathcal{U}^0$ is necessarily true, it follows that $\exists!kRk$. ⋈

(901.1) Assume $Op^1(H)$ and $Op^1(G)$. Then let R abbreviate $G \circ H$, where the latter was defined in (900.1) as $[\lambda xy\, \exists z(Hxz \,\&\, Gzy)]$. As noted in (900), then R exists, since its definiens is a core λ-expression and so exists by axiom (39.2). We have to show:

(A) $Op^1(R)$

(B) $\forall x(R(x) = G(H(x)))$

Proof of (A). By (894.1), we have to show:

(C) $Rigid(R)$

(D) $R : \mathbb{N} \longrightarrow \mathbb{N}$

To show (C), suppose not, for reductio. Then since R exists, it follows by the definition of rigidity (571.1) and modal reasoning, $\Diamond\exists x\exists y(Rxy \,\&\, \neg\Box Rxy)$. So, by CBF and further modal reasoning, $\exists x\exists y\Diamond(Rxy \,\&\, \Diamond\neg Rxy)$. Then let a and b be such objects so that we know $\Diamond(Rab \,\&\, \Diamond\neg Rab)$. The $\Diamond$ distributes over a conjunction but since $\Diamond\Diamond\varphi \to \Diamond\varphi$ (4$\Diamond$), it follows that:

(ϑ) $\Diamond Rab$

(ξ) $\Diamond\neg Rab$

Now we know, independently by β-Conversion and the existence of R, the following is a modally strict theorem:

(ζ) $Rab \equiv \exists z(Haz \,\&\, Gzb)$

From this and $\Diamond Rab$, it follows by a Rule of Substitution that $\Diamond\exists z(Haz \,\&\, Gzb)$. Hence, by BF$\Diamond$, $\exists z\Diamond(Haz \& Gzb)$. Let c be witness so that we know $\Diamond(Hac \& Gcb)$. Hence $\Diamond Hac \& \Diamond Gcb$, by (162.3). But since H and G are both operations, they are rigid (894.1), and so both $\Box(Hac \to \Box Hac)$ and $\Box(Gcb \to \Box Gcb)$. So it follows respectively by K$\Diamond$ that both that $\Diamond\Box Hac$ and $\Diamond\Box Gcb$. But then by (165.2), it follows respectively that $\Box Hac$ and $\Box Gcb$. Hence $\Box(Hac \& Gcb)$. So $\exists z\Box(Haz \& Gzb)$. By the Buridan formula (168.1), it follows that $\Box\exists z(Haz \,\&\, Gzb)$. So by the modally strict theorem (ζ) and a Rule of Substitution, it follows that $\Box Rab$, which contradicts (ξ).

To show (D), we have to show, by (874.2):

(E) $R \div \mathbb{N} \longrightarrow \mathbb{N}$

(F) $HasDomain(R, \mathbb{N})$

To show (E), we have to show, by (866.1):

(E.1) $R |: \mathbb{N} \longrightarrow \mathbb{N}$.

Proof. By (862), we have to show $\forall x(\mathbb{N}x \to \exists! y(\mathbb{N}y \,\&\, Rxy))$. So by GEN, assume $\mathbb{N}x$. Since both $Op^1(H)$ and $Op^1(G)$, let d be the unique number such that Hxd and let e be the unique number such that Gde. Since e is therefore the unique number such that $Hxd \,\&\, Gde$, it follows not only that Rxe, but also that e is the unique number such that Rxe. Thus e is the witness to $\exists! y(\mathbb{N}y \,\&\, Rxy)$

(E.2) $\forall x \forall y(\mathbb{N}x \,\&\, Rxy \to \mathbb{N}y)$.

Proof. By GEN, assume $\mathbb{N}x \,\&\, Rxy$. From the second conjunct, it follows by β-Conversion that $\exists z(Hxz \,\&\, Gzy)$. But since G is an operation, it maps any witness to this last claim to a number. So $\mathbb{N}y$.

To show (F), we have to show, by (872.1):

$\forall x(\mathbb{N}x \equiv \exists y Rxy)$

Proof. ($\to$) Assume $\mathbb{N}x$. But then $\exists y Rxy$ follows *a fortiori* from the reasoning used to establish (E.1). ($\leftarrow$) Assume $\exists y Rxy$. But then $\mathbb{N}x$ follows *a fortiori* from the reasoning used to establish (E.2).

Proof of (B). By Rule $\forall$I, it suffices to show, for an arbitrary number k, that $R(k) = G(H(k))$. Since we now know R is a unary operation, there is a unique j such that Rkj. So by β-conversion, we have $\exists z(Hkz \,\&\, Gzj)$. Let i be a witness, so $Hki \,\&\, Gij$. By (899.2) and (893), $i = H(k)$ and $j = G(i)$ and $j = R(k)$. By substitution of identicals, $j = G(H(k))$. So $R(k) = G(H(k))$.

(901.2) Assume $Op^n(H_1), \ldots, Op^n(H_m)$, and $Op^m(G)$. Then let R abbreviate the defined notation $G \circ (H_1, \ldots, H_m)$, where this was defined in (900.2) to be the relation:

$$[\lambda x_1 \ldots x_n y\, \exists z_1 \ldots \exists z_m(H_1 x_1 \ldots x_n z_1 \,\&\, \ldots \,\&\, H_m x_1 \ldots x_n z_m \,\&\, G z_1 \ldots z_m y)].$$

Then since the proof is a generalization of that used for the previous theorem, we leave it as an exercise. $\bowtie$

(905.1.b) By (898.4), we know $Op^1(\pi^2_1)$. So by the definition (.1.a), $Op^1(\boldsymbol{A}_0)$. $\bowtie$

(905.2.a) Assume $Op^1(\boldsymbol{A}_m)$. By (896), we know that $Op^1(\boldsymbol{s})$. Then by (901.1), that $Op^1(\boldsymbol{s} \circ \boldsymbol{A}_m)$. $\bowtie$

(905.3) We prove this by induction. The base case, $Op^1(\boldsymbol{A}_0)$ is established by (905.1.b). For the inductive case, our I.H. is that $Op^1(\boldsymbol{A}_m)$. Then, by (905.2.a), $Op^1(\boldsymbol{s} \circ \boldsymbol{A}_m)$. So by (905.2.b), $Op^1(\boldsymbol{A}_{m'})$. ⋈

(905.5) By the definition of a binary operation (894), we have to show:

(A) $Rigid(\boldsymbol{A})$

(B) $\boldsymbol{A} : \mathbb{N}^{\times 2} \longrightarrow \mathbb{N}$

(A) Suppose, for reductio, $\boldsymbol{A}$ isn't rigid. Then if we apply standard modal reasoning to the definition of rigidity, it follows that there are numbers a, b, c such that:

(ζ) $\Diamond \boldsymbol{A}abc \,\&\, \Diamond\neg\boldsymbol{A}abc$.

But now consider the fact that $Op^1\boldsymbol{A}_b$ (905.3), which implies $Rigid\boldsymbol{A}_b$ (894.1). Then by (899.3) there must be a unique number, say d, such that $\boldsymbol{A}_b ad$ and by (571.1) $\Box\boldsymbol{A}_b ad$. Hence, by definition (905.4), β-Conversion, and the fact that $n = m \equiv n \doteq m$, it follows that $\Box\boldsymbol{A}abd$. If $c = d$ we have a contradiction with the second conjunct of (ζ) and, otherwise, we have a contradiction with the first conjunct of (ζ).

(B) Suppose, for reductio, that it is not the case that $\boldsymbol{A} : \mathbb{N}^{\times 2} \longrightarrow \mathbb{N}$. Then, given that $\boldsymbol{A}$ takes only numbers as arguments (905.4), then one of the following must hold, by (875.2):

- there are numbers a and b for which $\boldsymbol{A}$ doesn't yield a unique number, or
- there are numbers a and b for which $\boldsymbol{A}$ yields something other than a number, or
- there are numbers a and b for which $\boldsymbol{A}$ yields nothing.

By (905.4), each of these cases implies that $\boldsymbol{A}_b(a)$ is not a number. But $Op^1(\boldsymbol{A}_b)$ (905.3) and (899.1) imply that $\boldsymbol{F}_b(a)$ is a number. ⋈

(905.6.a) This follows from (905.1.a) directly. ⋈

(905.6.b) The fact that $\boldsymbol{A}(n, m') = [\boldsymbol{s} \circ \boldsymbol{A}](n, m)$ follows from 905.2.b) via (905.4) directly. The result then follows immediately. ⋈

(908.1.b) (Exercise)

(908.2.a) Assume $Op^1(\boldsymbol{F}_m)$, where m is any natural number. By (898.4), we know that $Op^1(\pi^2_1)$. By (897.2), we know that $Op^1(\mathcal{C}^2_m)$. Then by (901.2), it follows that $Op^1(G \circ (\pi^2_1, \mathcal{C}^2_m, \boldsymbol{F}_m))$. ⋈

(908.3) We prove this by induction. The base case, $Op^1(\boldsymbol{F}_0)$ is established by (908.1.b). For the inductive case, our IH is that $Op^1(\boldsymbol{F}_m)$. Then, by (908.2.a), $Op^1(G \circ (\pi^2_1, \mathcal{C}^2_m, \boldsymbol{F}_m))$. So by definition of $\boldsymbol{F}_{m'}$ (908.2.b), $Op^1(\boldsymbol{F}_{m'})$. ⋈

(908.5) We prove our theorem by showing:

(A) $Rigid(\boldsymbol{F})$

(B) $\boldsymbol{F} : \mathbb{N}^{\times 2} \longrightarrow \mathbb{N}$.

(A) Suppose, for reductio, $\boldsymbol{F}$ isn't rigid. Then if we apply standard modal reasoning to the definition of rigidity, it follows that there are numbers a, b, c such that:

(ζ) $\Diamond \boldsymbol{F}abc \mathbin{\&} \Diamond\neg \boldsymbol{F}abc$.

But now consider the fact that $Op^1(\boldsymbol{F}_b)$ (908.3), which implies $Rigid(\boldsymbol{F}_b)$ (894.1). Then by (899.3) there must be a unique number, say d, such that $\boldsymbol{F}_b ad$ and by (571.1) $\Box \boldsymbol{F}_b ad$. Hence, by definition (908.4), β-Conversion, and the fact that $n = m \equiv n \doteq m$, it follows that $\Box \boldsymbol{F}abd$. If $c = d$ we have a contradiction with the second conjunct of (ζ) and, otherwise, we have a contradiction with the first conjunct of (ζ).

(B) Suppose, for reductio, that it is not the case that $\boldsymbol{F} : \mathbb{N}^{\times 2} \longrightarrow \mathbb{N}$. Then, given that $\boldsymbol{F}$ takes only numbers as arguments (908.4), one of the following must hold, by (875.2):

- there are numbers a and b for which $\boldsymbol{F}$ doesn't yield a unique number, or
- there are numbers a and b for which $\boldsymbol{F}$ yields something other than a number, or
- there are numbers a and b for which $\boldsymbol{F}$ yields nothing.

By (908.4), each of these cases implies that $\boldsymbol{F}_b(a)$ is not a number. But $Op^1(\boldsymbol{F}_b)$ (908.3) and (899.1) jointly imply that $\boldsymbol{F}_b(a)$ is a number. ⋈

(908.6.a) $\boldsymbol{F}(n,0) = \boldsymbol{F}_0(n) = H(n)$ by (908.4), β-Conversion, and (908.1.a). ⋈

(908.6.b) We may reason as follows:

$$\begin{array}{rcll}
\boldsymbol{F}(n,m') & = & \boldsymbol{F}_{m'}(n) & \text{by (908.4)} \\
 & = & [G \circ (\pi^2_1, \mathcal{C}^2_m, \boldsymbol{F}_m)](n) & \text{by (908.2.b)} \\
 & = & [G \circ (\pi^3_1, \pi^3_2, \boldsymbol{F})](n,m) & \text{by (898.2), (897.1), (908.4)} \\
 & = & G(n,m,\boldsymbol{F}(n,m)) & \text{by (898.2), (901.2)} \qquad ⋈
\end{array}$$

(909) By (908.5) and (908.6.a), and (908.6.b). ⋈

(910.1.b) (Exercise)

(910.2.a) Assume $Op^i(\boldsymbol{F}_m)$, where m is any natural number. By (898.4), we know that for all k, $Op^i(\pi^{i'}_k)$. By (897.2), we know that $Op^1(\mathcal{C}^{i'}_m)$. Then by

(901.2), it follows that $Op^{i''}(G \circ (\pi_1^{i'}, \ldots, \pi_i^{i'}, \mathcal{C}_m^{i'}, \boldsymbol{F}_m))$. Note that when $i = 0$, this still holds, though there are no projection functions in the composition. ⋈

(910.3) We prove this by induction. The base case, $Op^i(\boldsymbol{F}_0)$ is established by (910.1.b). For the inductive case, our IH is that $Op^i(\boldsymbol{F}_m)$. Then, by (910.2.a), $Op^i(G \circ (\pi_1^{i'}, \ldots, \pi_i^{i'}, \mathcal{C}_m^{i'}, \boldsymbol{F}_m))$. So by definition of $\boldsymbol{F}_{m'}$ (910.2.b), $Op^i(\boldsymbol{F}_{m'})$. ⋈

(910.5) We prove our theorem by showing:

(A) $Rigid(\boldsymbol{F})$

(B) $\boldsymbol{F} : \mathbb{N}^{\times(i+1)} \longrightarrow \mathbb{N}$.

(A) Suppose, for reductio, $\boldsymbol{F}$ isn't rigid. Then if we apply standard modal reasoning to the definition of rigidity, it follows that there are numbers $a_1, \ldots, a_i$, b, and c such that:

(ζ) $\Diamond \boldsymbol{F} a_1 \ldots a_i bc \;\&\; \Diamond \neg \boldsymbol{F} a_1 \ldots a_i bc$.

But now consider the fact that $Op^i(\boldsymbol{F}_b)$ (910.3), which implies $Rigid(\boldsymbol{F}_b)$ (894.1). Then by (899.3) there must be a unique number, say d, such that $\boldsymbol{F}_b a_1 \ldots a_i d$ and by (571.1) $\Box \boldsymbol{F}_b a_1 \ldots a_i d$. Hence, by definition (910.4), β-Conversion, and the fact that $n = m \equiv n \doteq m$, it follows that $\Box \boldsymbol{F} a_1 \ldots a_i bd$. If $c = d$ we have a contradiction with the second conjunct of (ζ) and, otherwise, we have a contradiction with the first conjunct of (ζ). Note that when $i = 0$, this still holds, though no numbers $a_1, \ldots, a_i$ are used.

(B) Suppose, for reductio, that it is not the case that $\boldsymbol{F} : \mathbb{N}^{\times(i+1)} \longrightarrow \mathbb{N}$. Then, given that $\boldsymbol{F}$ takes only numbers as arguments (910.4), one of the following must hold, by (875.2):

- there are numbers $a_1, \ldots, a_i$, and b for which $\boldsymbol{F}$ doesn't yield a unique number, or
- there are numbers $a_1, \ldots, a_i$, and b for which $\boldsymbol{F}$ yields something other than a number, or
- there are numbers $a_1, \ldots, a_i$, and b for which $\boldsymbol{F}$ yields nothing.

By (910.4), each of these cases implies that $\boldsymbol{F}_b(a_1, \ldots, a_i)$ is not a number. But this contradicts $Op^i(\boldsymbol{F}_b)$ (910.3). Note that when $i = 0$, this still holds, though no numbers $a_1, \ldots, a_i$ are used and $\boldsymbol{F}_b$ is nullary. ⋈

(910.6.a) $\boldsymbol{F}(n_1, \ldots, n_i, 0) = \boldsymbol{F}_0(n_1, \ldots, n_i) = H(n_1, \ldots, n_i)$, by (910.4), β-Conversion, and (910.1.a). ⋈

(910.6.b) We may reason as follows, where $i \geq 0$:

$$\begin{aligned}
\boldsymbol{F}(n_1,\ldots,n_i,m') &= \boldsymbol{F}_{m'}(n_1,\ldots,n_i) && (910.4)\\
&= [G\circ(\pi_1^{i'},\ldots,\pi_i^{i'},\mathcal{C}_{m'}^{i'},\boldsymbol{F}_m)](n_1,\ldots,n_i) && (910.2.b)\\
&= [G\circ(\pi_1^{i''},\ldots,\pi_i^{i''},\pi_{i+1}^{i''},\boldsymbol{F})](n_1,\ldots,n_i,m) && (898.2), (897.1), (910.4)\\
&= G(n_1,\ldots,n_i,m,\boldsymbol{F}(n_1,\ldots,n_i,m)) && (898.2), (901.2)
\end{aligned}$$

Note that when $i = 0$, this still holds, though no variables $n_1,\ldots,n_i$ are used and $\boldsymbol{F}_{m'}$ is nullary. ⋈

(911) Assume $Op^i(H)\,\&\,Op^{i''}(G))$, where $i \geq 0$. Then it suffices to show that $\boldsymbol{F}$, as defined in (910.4), is a witness to:

$$\exists F(Op^{i'}(F)\,\&\,F(n_1,\ldots,n_i,0) = H(n_1,\ldots,n_i)\,\&\\ F(n_1,\ldots,n_i,m') = G(n_1,\ldots,n_i,m,F(n_1,\ldots,n_i,m)))$$

By (910.5), it follows that $Op^{i'}(\boldsymbol{F})$. By (910.6.a), $\boldsymbol{F}(n_1,\ldots,n_i,0) = H(n_1,\ldots,n_i)$. And by (910.6.b), $\boldsymbol{F}(n_1,\ldots,n_i,m') = G(n_1,\ldots,n_i,m,\boldsymbol{F}(n_1,\ldots,n_i,m))$. Note that when $i = 0$, this still holds, though no variables $n_1,\ldots,n_i$ are used and H is nullary. ⋈

(912.1.b) Note that in the argument below for the recursive clause, $\boldsymbol{s}\circ\pi_3^4$ only makes use of its third argument:

Base clause:

$\boldsymbol{A}(n,0)$	=	$\pi_1^2(n)$	by (.1.a) and (910.6.a)
	=	n	by definition of π_1^2 in (898.1)

Recursive Clause:

$\boldsymbol{A}(n,m')$	=	$[\boldsymbol{s}\circ\pi_3^4](n,m,\boldsymbol{A}(n,m))$	by (.1.a) and (910.6.b)
	=	$(\boldsymbol{A}(n,m))'$	by the definitions of π_3^4 (898.3) ⋈

(912.2.b) The argument is easy:

Base clause:

$\boldsymbol{M}(n,0)$	=	$\mathcal{C}_0^2(n)$	by (.2.a) and (910.6.a)
	=	0	by definition of $\mathcal{C}_0^2$ in (897.1)

Recursive Clause:

$\boldsymbol{M}(n,m')$	=	$[\boldsymbol{A}\circ(\pi_1^4,\pi_3^4)](n,m,\boldsymbol{M}(n,m))$	by (.2.a) and (910.6.b)
	=	$n+\boldsymbol{M}(n,m)$	by definitions π_1^4, π_3^4, $\boldsymbol{A}$, and $\circ$ ⋈

(912.3.b) The argument is easy:

Base clause:

$\boldsymbol{E}(n,0)$	=	$\mathcal{C}_1^2(n)$	by (.3.a) and (910.6.a)
	=	1	by definition of $\mathcal{C}_1^2$ in (897.1)

Recursive Clause:

$\boldsymbol{E}(n,m')$	=	$[\boldsymbol{M}\circ(\pi_1^4,\pi_3^4)](n,m,\boldsymbol{E}(n,m))$	by (.3.a) and (910.6.b)
	=	$n\times\boldsymbol{E}(n,m)$	by dfs of π_1^4, π_3^4, $\boldsymbol{M}$ and $\circ$ ⋈

(912.4.b) The argument is:

Base clause:

$\mathbf{\Pi}(0)$	$=$	$\mathcal{C}^1_1$	by (.4.a) and (910.6.a)
	$=$	1	by definition of $\mathcal{C}^1_1$ in (897.1)

Recursive Clause:

$\mathbf{\Pi}(n')$	$=$	$[\boldsymbol{M} \circ (\pi^3_2, \boldsymbol{s} \circ \pi^3_1)](n, \mathbf{\Pi}(n))$	by (.4.a) and (910.6.b)
	$=$	$\mathbf{\Pi}(n) \times n'$	by dfs of $\boldsymbol{M}$, π^3_2, $\boldsymbol{s}$, π^3_1, and $\circ \bowtie$

(916) We show each axiom in turn is a theorem.

• $n' \neq 0$. By theorem (810) and (821).

• $n' = m' \to n = m$. By theorem (811) and (821).

• $n + 0 = n$. By theorem (912.1.b).

• $n + m' = (n + m)'$. By theorem (912.1.b).

• $n \times 0 = 0$. By theorem (912.2.b).

• $n \times m' = n + (n \times m)$. By theorem (912.2.b).

• $\neg(n < 0)$. By theorems (805.2) and (832.1).

• $n < m' \equiv (n < m \lor n = m)$. $(\to)$ Assume $n < m'$. By (796.2), we know $(\underline{G}xy \,\&\, \underline{G}^*zy) \to \underline{G}^+zx$. Substituting $\mathbb{P}$ for $\underline{G}$, m for x, m' for y, and n for z, we obtain: $(\mathbb{P}mm' \,\&\, \mathbb{P}^*nm') \to \mathbb{P}^+nm$. Since $\mathbb{P}mm'$ is a theorem, and our assumption $n < m'$ implies $\mathbb{P}^*nm'$ (832.1), it follows that $\mathbb{P}^+nm$, i.e., by (832.2), $n \leq m$, i.e., by (840.7), $n < m \lor n = m$. $(\leftarrow)$ (Exercise)

• Induction Axiom:

$$(F0 \,\&\, \forall n(Fn \to Fn')) \to \forall nFn$$

By theorems (812) and (821).

• Comprehension Scheme:

$$\exists F \forall n(Fn \equiv \varphi^*),$$
where φ^* is the translation of φ and F doesn't occur free in φ^*

Proof. It suffices to show that $[\lambda n\, \varphi^*]$ exists and is a witness to the existential claim. Note that we can expand the restricted variable by choosing an unrestricted variable, say y, that doesn't occur free in $\varphi^*(n)$. Then it suffices to show that $[\lambda y\, \mathbb{N}y \,\&\, \varphi^*]\!\downarrow$ and is a witness to the existential claim. But since there are no encoding formulas in the translations of the atomic formulas of 2nd order PA, and the translations of the complex formulas don't introduce any encoding formulas, no variables occur in encoding position in φ^* (9.1). Hence, $[\lambda y\, \mathbb{N}y \,\&\, \varphi^*]$ is a core λ-expression (9.2) and so $[\lambda y\, \mathbb{N}y \,\&\, \varphi^*]\!\downarrow$ by (39.2). So by GEN, it remains to show $[\lambda y\, \mathbb{N}y \,\&\, \varphi^*]n \equiv \varphi^*$, which we do by the following biconditional chain:

$[\lambda y\, \mathbb{N}y \,\&\, \varphi^*]n$

$\equiv$	$n{=}n \,\&\, [\lambda y\, \mathbb{N}y \,\&\, \varphi^*]n$	$(\rightarrow)$ =I, &I / $(\leftarrow)$ &E
$\equiv$	$x{=}n \,\&\, [\lambda y\, \mathbb{N}y \,\&\, \varphi^*]x$	=I / =E, $(\rightarrow)$ uses fresh variable x
$\equiv$	$x{=}n \,\&\, (\mathbb{N}x \rightarrow \varphi^{*x}_{\ n})$	β-Conversion with unrestricted variable
$\equiv$	$x{=}n \,\&\, \mathbb{N}x \,\&\, (\mathbb{N}x \rightarrow \varphi^{*x}_{\ n})$	since $x{=}n$, $\mathbb{N}x$
$\equiv$	$x{=}n \,\&\, \varphi^{*x}_{\ n}$	$(\rightarrow)$ MP / $(\leftarrow)$ since $x{=}n$, $\mathbb{N}x$ and (38.1)
$\equiv$	$n{=}n \,\&\, \varphi^*$	=E / =I, $(\leftarrow)$ use with fresh variable x
$\equiv$	φ^*	$(\rightarrow)$ &E / $(\leftarrow)$ =I and &I $\bowtie$

(918.1) We proceed by induction on m. **Base Case:** $m = 0$. By GEN, it suffices to prove $\mathbb{P}^+xm \rightarrow \mathbb{N}x$. So assume $\mathbb{P}^+xm$. Then $\mathbb{P}^+x0$. Since we know, by (805.2), that $\neg\mathbb{P}^*x0$, it follows by a fact about $\mathbb{P}^+$ (806.2) that $x = 0$. Since $\mathbb{N}0$ (808), it follows that $\mathbb{N}x$.

Inductive Case: Assume our theorem holds for m, so that our inductive hypothesis is:

(IH) $\forall x(\mathbb{P}^+xm \rightarrow \mathbb{N}x)$

Then we need to show: $\forall x(\mathbb{P}^+xm' \rightarrow \mathbb{N}x)$. So, by GEN, assume $\mathbb{P}^+xm'$. Then by a fact about $\mathbb{P}^+$ (806.2) either $\mathbb{P}^*xm' \vee x = m'$. We reason to the conclusion $\mathbb{N}x$ by cases from the disjuncts.

- Suppose $\mathbb{P}^*xm'$. Note that if we apply GEN twice to the modally strict theorem (803.5) and then RN to the result, we obtain:

  $\Box\forall x\forall y(\mathbb{P}xy \rightarrow (D!x \,\&\, D!y))$

  From this and the fact that $\mathbb{P}\!\downarrow$ (801.2), it follows by (791.1) that $\mathbb{P}$ is a relation on discernibles. So the following is an instance of (796.2):

  (ϑ) $\mathit{1\text{-}1}(\mathbb{P}) \rightarrow ((\mathbb{P}mm' \,\&\, \mathbb{P}^*xm') \rightarrow \mathbb{P}^+xm)$

  But by (802.3), $\mathbb{P}$ is a 1-1 relation. Moreover, by (824), $\mathbb{P}mm'$. And since we're working under the hypothesis that $\mathbb{P}^*xm'$, we have everything we need to infer from (ϑ) that $\mathbb{P}^+xm$. So by our IH, $\mathbb{N}x$.

- Suppose $x = m'$. Then since $\mathbb{N}m'$ by hypothesis, we have $\mathbb{N}x$. $\bowtie$

(918.2) Suppose, for reductio, $\exists n Numbers(n, \mathbb{N})$, and let a be a witness, so that we know $\mathbb{N}a$ and $Numbers(a, \mathbb{N})$. Then by (817.6), it follows that:

$\exists y(Numbers(y, [\lambda z\, \mathbb{P}^+za]) \,\&\, \mathbb{P}ay)$

Suppose b is a witness, so that we have $Numbers(b,[\lambda z\, \mathbb{P}^{+}za])\,\&\,\mathbb{P}ab$. From the second conjunct and $\mathbb{N}a$, it follows that $\mathbb{N}b$, by (814.1). It also follows, from the second conjunct and the fact that a and b are natural numbers, that $a < b$, by (833.1). Now, independently, in theorem (835), instantiate $[\lambda z\, \mathbb{P}^{+}za]$ for F, $\mathbb{N}$ for G, a for m, and b for n, so that we know:

$$(\vartheta)\quad \big(Numbers(b,[\lambda z\, \mathbb{P}^{+}za])\,\&\,Numbers(a,\mathbb{N})\,\&\,\forall u([\lambda z\, \mathbb{P}^{+}za]u \rightarrow \mathbb{N}u)\big) \rightarrow b \leq a$$

We already know the first two conjuncts of the antecedent of (ϑ). To establish the third conjunct, note that since (918.1) holds for any natural number m, it holds for a. Hence $\forall x(\mathbb{P}^{+}xa \rightarrow \mathbb{N}x)$. Since $[\lambda z\, \mathbb{P}^{+}za]\downarrow$, it follows, by an instance of λ-Conversion and a Rule of Substitution, that $\forall x([\lambda z\, \mathbb{P}^{+}za]x \rightarrow \mathbb{N}x)$. The third conjunct of the antecedent of (ϑ) follows from this *a fortiori*. Hence $b \leq a$. But we previously established $a < b$. So by (833.7), $a < a$, which contradicts (838.1) $\bowtie$

(918.3) By definitions (917.2) and (917.1), we have to show $\neg\mathbb{N}\#\mathbb{N}$. Note first that by (809.2), $Rigid(\mathbb{N})$. So by a modally strict fact about numbering and rigid properties (774.5), $Numbers(\#\mathbb{N},\mathbb{N})$. Now assume, for reductio, that $\mathbb{N}\#\mathbb{N}$. Then $\exists n Numbers(n,\mathbb{N})$, which contradicts (918.2). $\bowtie$

(918.4) By (777). $\bowtie$

(918.5) (Exercise)

(920.1) – **(920.5)** (Exercises)

(920.6) For reductio, assume some natural number, say n, is such that $\mathbb{P}n\aleph_0$. Then by (814.1), $\mathbb{N}\aleph_0$. But this contradicts (920.4). $\bowtie$

(920.7) Since $\mathbb{N}$ is rigid (809.2), it follows by (774.5) that $Numbers(\#\mathbb{N},\mathbb{N})$. But $\aleph_0 = \#\mathbb{N}$, by definition (919). So $Numbers(\aleph_0,\mathbb{N})$ and, so by (778.1), $\aleph_0$ is a natural cardinal. Now since 0 is discernible, consider $\mathbb{N}^{-0}$. Since we know that something numbers this property (763.1), suppose a is such that $Numbers(a,\mathbb{N}^{-0})$. Then, similarly, a is a natural cardinal (778.1). Hence, where $\mathbb{N}$ and 0 are witnesses, we have:

$$\exists F\exists u(Fu\,\&\,Numbers(\aleph_0,F)\,\&\,Numbers(a,F^{-u}))$$

So it follows by (801.3) that:

$$(\vartheta)\quad \mathbb{P}a\aleph_0$$

Note independently, that as an instance of (764.1), we know:

$$\mathbb{N} \approx_D \mathbb{N}^{-0} \rightarrow (Numbers(\aleph_0,\mathbb{N}) \equiv Numbers(\aleph_0,\mathbb{N}^{-0}))$$

But with a little work, one can show $\mathbb{N} \approx_D \mathbb{N}^{-0}$:

Proof. We establish this from the definition of $\approx_D$ (747.3) by showing that the successor function s is the witness, i.e., that $s\,|\!:\mathbb{N} \xleftrightarrow{\text{1-1}}_D \mathbb{N}^{-0}$, i.e., that s correlates$_D$ $\mathbb{N}$ and $\mathbb{N}^{-0}$ one-to-one (747.2). Note first that s is 1-1, given that $\mathbb{P}$ is 1-1 (802.3) and the definition of s as the predecessor relation restricted to the numbers (879.2). Moreover, s has $\mathbb{N}_+$ as its range (879.1). So if we can show that $\forall x(\mathbb{N}_+x \equiv \mathbb{N}^{-0}x)$, then it follows that s has $\mathbb{N}^{-0}$ as its range (872.2). By GEN, it remains to show $\mathbb{N}_+x \equiv \mathbb{N}^{-0}x$. By definitions (752.2), (841.1), and β-Conversion, we have to show:

$$x > 0 \equiv (\mathbb{N}x \,\&\, x \neq 0)$$

We prove the commuted form, $(\mathbb{N}x \,\&\, x \neq 0) \equiv x > 0$, as follows:

$$\begin{array}{ll}
\multicolumn{2}{l}{\mathbb{N}x \,\&\, x \neq 0} \\
\quad \equiv \mathbb{N}x \,\&\, \mathbb{N}x \,\&\, x \neq 0 & \text{repeat/eliminate } \mathbb{N}x \text{ conjunct} \\
\quad \equiv \mathbb{N}x \,\&\, (\mathbb{P}^{+}0x \,\&\, x \neq 0) & \text{expand/apply df. } \mathbb{N}\text{, group} \\
\quad \equiv \mathbb{N}x \,\&\, ((\mathbb{P}^{*}0x \vee x = 0) \,\&\, x \neq 0) & \text{expand/apply definition } \mathbb{P}^{+} \\
\quad \equiv \mathbb{N}x \,\&\, ((\mathbb{P}^{*}0x \,\&\, x \neq 0) \vee (x = 0 \,\&\, x \neq 0)) & \text{distribute/contract } \vee, \& \\
\quad \equiv \mathbb{N}x \,\&\, (\mathbb{P}^{*}0x \,\&\, x \neq 0) & \text{eliminate/add impossibility} \\
\quad \equiv \mathbb{N}x \,\&\, \mathbb{P}^{*}0x & \text{simplify/apply } \neg\mathbb{P}^{*}00 \text{ (805.3)} \\
\quad \equiv [\lambda x\, \mathbb{N}x \,\&\, \mathbb{P}^{*}0x]x & \beta\text{-Conversion} \\
\quad \equiv 0 < x & \text{apply/expand, (828.1), (830.1)} \\
\quad \equiv x > 0 & \text{(832.3)}
\end{array}$$

Hence, $\mathit{Numbers}(\aleph_0, \mathbb{N}) \equiv \mathit{Numbers}(\aleph_0, \mathbb{N}^{-0})$. Since we already know the left side, it follows that $\mathit{Numbers}(\aleph_0, \mathbb{N}^{-0})$. But by (763.2), there is a unique number that numbers $\mathbb{N}^{-0}$. Hence $a = \aleph_0$. Then by (ϑ), $\mathbb{P}\aleph_0\aleph_0$. ⋈

(923) By definition (922.2) and a Rule of Substitution, we have to show:

$$\exists x \exists G(\mathit{InfiniteClassOf}(x, G))$$

So by definition (922.2) and a Rule of Substitution, we have to show:

$$(\vartheta)\ \ \exists x \exists G(G{\downarrow} \,\&\, \mathit{ClassOf}(x, G) \,\&\, \exists \kappa(\mathit{Infinite}(\kappa) \,\&\, \mathit{Numbers}(\kappa, G)))$$

But consider $\mathbb{N}$. By (315.1), $\exists x \mathit{ClassOf}(x, \mathbb{N})$. Suppose a is such an object, so that we know $\mathit{ClassOf}(a, \mathbb{N})$. Then we may prove our theorem by showing that a and $\mathbb{N}$ are witnesses to (ϑ). And given what we've established thus far and the fact that $\#\mathbb{N}$ is a natural cardinal (918.4), it remains, by &I and ∃I, only to show $\mathit{Infinite}(\#\mathbb{N})$ and $\mathit{Numbers}(\#\mathbb{N}, \mathbb{N})$. But the first just is (918.3) and the second follows from (774.5) and the rigidity of $\mathbb{N}$ (809.2). ⋈

(937.4) (Exercise)

(937.6) – **(937.7)** (Exercises)

(938) – (939) (Exercises)

(940.1) Let φ be any formula other than a description of type $\langle\rangle$. Then φ is either a constant or variable of type $\langle\rangle$ or a non-basic formula. So we establish our theorem by cases:

> *Case* (1): φ is a constant or variable of type $\langle\rangle$. Then $\varphi\!\downarrow$ by (935.5).
>
> *Case* (2): φ is a non-basic formula. Then by (935.23), $O!\varphi$. Hence, by (935.8.a), $\varphi\!\downarrow$. ⋈

(940.2) By the previous theorem (940.1), $\varphi\!\downarrow$, for any formula φ other than a description of type $\langle\rangle$. But by the BNF definition in (928), these formulas are all the relation terms Π of type $\langle\rangle$. So $\Pi^{\langle\rangle}\!\downarrow$, for any relation term of type $\langle\rangle$ other than a description of type $\langle\rangle$. ⋈

(942) By (935.5), the modal closures of $\alpha\!\downarrow$ are axioms, where α is a variable of any type. So $\Box\alpha\!\downarrow$ is an axiom. Hence, by GEN, $\forall\alpha\Box\alpha\!\downarrow$. Now let φ be $\Box\alpha\!\downarrow$. Then, where τ is a term having the same type as α, we have the following instance of axiom (935.4):

> $\forall\alpha\Box\alpha\!\downarrow\,\rightarrow(\tau\!\downarrow\,\rightarrow\Box\tau\!\downarrow)$, where τ is *any* term substitutable for α in $\Box\alpha\!\downarrow$

But every term τ of the same type as α is substitutable for α in $\Box\alpha\!\downarrow$. So it follows by Rule MP that $\tau\!\downarrow\,\rightarrow\Box\tau\!\downarrow$. ⋈

(943.1) We reason by cases from the definition of = in (933.9) – (933.12).

Case 1. Let τ and σ be arbitrary terms of type i, and assume $\tau=\sigma$. Then, where $O!$, $A!$, and F have type $\langle i\rangle$, it follows by definition (933.9) that:

$$(O!\tau \,\&\, O!\sigma \,\&\, \Box\forall F(F\tau \equiv F\sigma)) \vee (A!\tau \,\&\, A!\sigma \,\&\, \Box\forall F(\tau F \equiv \sigma F))$$

Clearly both disjuncts contain a conjunct ($O!\tau$ in the first disjunct and $A!\tau$ in the second) that implies, by (935.8), that $\tau\!\downarrow$.

Case 2. Let τ and σ be arbitrary terms, say Π and Π', of type $\langle t\rangle$, where t is any type, and assume $\tau=\sigma$, i.e., $\Pi=\Pi'$. Then, where x has type t, and $O!$, $A!$, and $\mathcal{H}$ have type $\langle\langle t\rangle\rangle$, it follows by definition (933.10) that:

$$(O!\Pi \,\&\, O!\Pi' \,\&\, \Box\forall x(x\Pi \equiv x\Pi')) \vee (A!\Pi \,\&\, A!\Pi' \,\&\, \Box\forall\mathcal{H}(\Pi\mathcal{H} \equiv \Pi'\mathcal{H}))$$

Again both disjuncts contain a conjunct ($O!\Pi$ in the first disjunct and $A!\Pi$ in the second) that implies, by (935.8), that $\Pi\!\downarrow$, i.e., $\tau\!\downarrow$.

Case 3. By reasoning analogous to Case 2.

Case 4. By reasoning analogous to Case 2. ⋈

(943.2) (Exercise)

(944.1) Let Π and Π' be any relation terms of type $\langle t_1,\ldots,t_n\rangle$ ($n \geq 0$), where $x_1,\ldots,x_n$ are variables of type $t_1,\ldots,t_n$, respectively, that don't occur free in Π and Π'. Assume $\Pi = \Pi'$. Then by (943.1), it follows that $\Pi\downarrow$, and by (943.2), it follows that $\Pi'\downarrow$. Note independently that where F and G are variables of the same type as Π and Π', respectively, then the following is axiomatic, since it is a closure of the axiom for the substitution of identicals (935.9):

$$\forall F\forall G(F=G \to (\Box\forall x_1\ldots x_n(Fx_1\ldots x_n \equiv Fx_1\ldots x_n) \to \Box\forall x_1\ldots x_n(Fx_1\ldots x_n \equiv Gx_x\ldots x_n)))$$

Since $x_1,\ldots,x_n$ don't occur free in $\Pi\downarrow$ and $\Pi'\downarrow$, they are substitutable, respectively, for F and G in the matrix of the above universal claim. So it follows by Rule $\forall$E (939) [93.1] that:

$$\begin{aligned}&\Pi=\Pi' \to\\ &\quad(\Box\forall x_1\ldots x_n(\Pi x_1\ldots x_n \equiv \Pi x_1\ldots x_n) \to \Box\forall x_1\ldots x_n(\Pi x_1\ldots x_n \equiv \Pi' x_x\ldots x_n))\end{aligned}$$

And since $\Pi = \Pi'$ by assumption, it follows that:

$$\Box\forall x_1\ldots x_n(\Pi x_1\ldots x_n \equiv \Pi x_1\ldots x_n) \to \Box\forall x_1\ldots x_n(\Pi x_1\ldots x_n \equiv \Pi' x_x\ldots x_n)$$

But the antecedent is clearly a theorem, by n applications of GEN and an application of RN to the instance $\Pi x_1\ldots x_n \equiv \Pi x_1\ldots x_n$ of the tautology $\varphi \equiv \varphi$. Hence $\Box\forall x_1\ldots x_n(\Pi x_1\ldots x_n \equiv \Pi' x_x\ldots x_n)$. ⋈

(944.2) (Exercise)

(945) The proof is analogous to the proof of (110), but appeals to: (a) the theorems in (937) [63] instead of the theorems in (63), (b) the theorems in (943) instead of (107), (c) the typed axiom for the substitution of identicals (935.9) instead of the second-order version (41), and (d) Rule $\forall$E (939) [93] instead of to (93). ⋈

(946.1) Let t be any type and suppose x is a variable of type t. Then by axiom (935.5), $[\lambda x\, \Diamond E!x]\downarrow$. Now the definition of $O!$ (933.7) and the Rule of Definition by Identity (937) [73] imply:

$$([\lambda x\, \Diamond E!x]\downarrow \to (O! = [\lambda x\, \Diamond E!x]))\ \&\ (\neg[\lambda x\, \Diamond E!x]\downarrow \to \neg O!\downarrow)$$

Hence $O! = [\lambda x\, \Diamond E!x]$, and so by (943.1), $O!\downarrow$. ⋈

(946.2) (Exercise)

(946.3) Let t be any type, x be a variable of type t, and $E!$, $O!$, and $A!$ have type $\langle t\rangle$. We want to prove $O!x \vee A!x$. So for reductio, assume $\neg(O!x \vee A!x)$. Then $\neg O!x$ and $\neg A!x$. But since $[\lambda x\, \Diamond E!x]\downarrow$ and $[\lambda x\, \neg\Diamond E!x]\downarrow$, we independently know, by the definitions of $O!$ and $A!$ and our Rule of Definition by Identity, that:

$$O! = [\lambda x\, \Diamond E!x]$$

$$A! = [\lambda x\, \neg\Diamond E!x]$$

Hence by Rule =E:

$$\neg[\lambda x\, \Diamond E!x]x$$

$$\neg[\lambda x\, \neg\Diamond E!x]x$$

But it also follows from $[\lambda x\, \Diamond E!x]\downarrow$ and $[\lambda x\, \neg\Diamond E!x]\downarrow$, that the following equivalences hold, by β-Conversion (935.26):

$$[\lambda x\, \Diamond E!x]x \equiv \Diamond E!x$$

$$[\lambda x\, \neg\Diamond E!x]x \equiv \neg\Diamond E!x$$

Hence we can conclude:

$$\neg\Diamond E!x$$

$$\neg\neg\Diamond E!x$$

But the latter implies $\Diamond E!x$. Contradiction. ⋈

(947.1) Given the four cases of the definition of identity (933.9) – (933.12), we may prove our theorem by establishing the following four cases:

Case 1: $x = x$, where x is a variable of type i.

Case 2: $F = F$, where F is a variable having a type of the form $\langle t\rangle$, for some type t.

Case 3: $F = F$, where F is a variable having a type of the form $\langle t_1, \ldots, t_n\rangle$ $(n \geq 2)$, for some types $t_1, \ldots, t_n$.

Case 4: $p = p$, where p is a variable of type $\langle\,\rangle$.

Case 1: Where x has type i, and where $O!$, $A!$, and F have type $\langle i\rangle$, we have to show, by definition (933.9):

$$O!x \,\&\, O!x \,\&\, \Box\forall F(Fx \equiv Fx)) \vee (A!x \,\&\, A!x \,\&\, \Box\forall F(xF \equiv xF))$$

The proof proceeds by disjunctive syllogism from the fact $O!x \vee A!x$ (946.3). (Exercise)

Case 2: Then F has type $\langle t\rangle$, for some type t, and so where x has type t, and $O!$, $A!$, and F have type $\langle\langle t\rangle\rangle$, we have to show, by definition (933.10):

$$(O!\mathit{F} \,\&\, O!\mathit{F} \,\&\, \Box\forall x(xF \equiv xF)) \vee (A!\mathit{F} \,\&\, A!\mathit{F} \,\&\, \Box\forall\mathcal{H}(\mathit{F}\mathcal{H} \equiv \mathit{F}\mathcal{H}))$$

The proof proceeds by disjunctive syllogism from the fact $O!F \vee A!F$ (946.3). (Exercise)

Case 3: Then F has type $\langle t_1,\ldots,t_n\rangle$ ($n \geq 2$), for some types $t_1,\ldots,t_n$, and so where $x_1,\ldots,x_n$ have types $t_1,\ldots,t_n$, respectively, and $O!$, $A!$, and F have type $\langle\langle t_1,\ldots,t_n\rangle\rangle$, we have to show, by definition (933.11):

$$\begin{array}{l} O!F \,\&\, O!F \,\&\, \forall x_2 \ldots \forall x_n([\lambda x_1\, Fx_1 \ldots x_n] = [\lambda x_1\, Fx_1 \ldots x_n]) \,\&\, \\ \forall x_1 \forall x_3 \ldots \forall x_n([\lambda x_2\, Fx_1 \ldots x_n] = [\lambda x_2\, Fx_1 \ldots x_n]) \,\&\ldots\&\, \\ \forall x_1 \ldots \forall x_{n-1}([\lambda x_n\, Fx_1 \ldots x_n] = [\lambda x_n\, Fx_1 \ldots x_n]) \quad \vee \\ \\ A!F \,\&\, A!F \,\&\, \Box\forall\mathcal{H}(F\mathcal{H} \equiv F\mathcal{H}) \end{array}$$

The proof proceeds by disjunctive syllogism from the fact $O!F \vee A!F$ (946.3). (Exercise)

Case 4: Then p has type $\langle\rangle$, and so where x has type i, and $O!$, $A!$, and F have type $\langle\rangle$, we have to show, by definition (933.12):

$$(O!p \,\&\, O!q \,\&\, [\lambda x\, p] = [\lambda x\, q]) \vee (A!p \,\&\, A!q \,\&\, \Box\forall\mathcal{H}(p\mathcal{H} \equiv q\mathcal{H}))$$

The proof proceeds by disjunctive syllogism from the fact that $O!p \vee A!p$ (946.3). (Exercise) ⋈

(947.2) Let t be any type, and let x and y be variables of type t. Assume the antecedent, $x = y$. Now by (117.1), we know $x = x$. Hence by the Variant version of Rule =E (945), it follows that $y = x$. ⋈

(947.3) Let t be any type, and let x, y, and z be variables of type t. Now assume the antecedent, so that by &E we know both $x = y$ and $y = z$. Then by Rule =E (110) it follows that $x = z$. ⋈

(947.4) Let t be any type and x, y, and z all be distinct variables of type t. Then use reasoning analogous to that in (117.4). ⋈

(948.1) – **(948.2)** (Exercises)

(949.1) Let $t_1,\ldots,t_n$ be any types. Then, by hypothesis:

(A) $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$ is a definition-by-= in which the variables $\alpha_1,\ldots,\alpha_n$ occur free ($n \geq 0$) and have types $t_1,\ldots,t_n$, respectively,

(B) $\tau_1,\ldots,\tau_n$ are any terms substitutable, respectively, for $\alpha_1,\ldots,\alpha_n$ in both definiens and definiendum, and

(C) $\Gamma \vdash \sigma(\tau_1,\ldots,\tau_n)\downarrow$.

From (A), (B), and the Rule of Definition by Identity (937.9) [73], we know:

$$\vdash (\sigma(\tau_1,\ldots,\tau_n)\downarrow \rightarrow \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n)) \,\&\, (\neg\sigma(\tau_1,\ldots,\tau_n)\downarrow \rightarrow \neg\tau(\tau_1,\ldots,\tau_n)\downarrow)$$

By (937.4) [63.3], the above holds for any premise set Γ:

(D) $\Gamma \vdash (\sigma(\tau_1,\ldots,\tau_n)\downarrow \to \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n))\ \&\ (\neg\sigma(\tau_1,\ldots,\tau_n)\downarrow \to \neg\tau(\tau_1,\ldots,\tau_n)\downarrow)$

By Rule &E (938) [86.2.a], it follows from (D) that:

(E) $\Gamma \vdash \sigma(\tau_1,\ldots,\tau_n)\downarrow) \ \to \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n)$

Hence, from (E) and (C) it follows by (937.4) [63.5] that:

$$\Gamma \vdash \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n) \qquad \bowtie$$

(949.2.a) By hypothesis:

(A) $\tau(\alpha_1,\ldots,\alpha_n) =_{df} \sigma(\alpha_1,\ldots,\alpha_n)$ is a definition-by-= in which the variables $\alpha_1,\ldots,\alpha_n$ occur free ($n \geq 0$) and have types $t_1,\ldots,t_n$, respectively,

(B) $\tau_1,\ldots,\tau_n$ are substitutable for $\alpha_1,\ldots,\alpha_n$, respectively, in both definiens and definiendum,

(C) φ contains one or more occurrences of $\tau(\tau_1,\ldots,\tau_n)$, and

(D) φ' is the result of replacing zero or more occurrences of $\tau(\tau_1,\ldots,\tau_n)$ in φ by $\sigma(\tau_1,\ldots,\tau_n)$

Now assume:

(E) $\Gamma \vdash \sigma(\tau_1,\ldots,\tau_n)\downarrow$

(F) $\Gamma \vdash \varphi$

(A), (B), and (E) imply, by the Rule of Identity by Definition (937) [120.1], that:

(G) $\Gamma \vdash \tau(\tau_1,\ldots,\tau_n) = \sigma(\tau_1,\ldots,\tau_n)$

Then by (C), (D), and Rule =E (945), it follows from (G) and (F) that $\Gamma \vdash \varphi'$. $\bowtie$

(949.2.b) (Exercise)

(950.1) Assume $O!\varphi$, where φ is any formula. Then $\varphi\downarrow$, by axiom (935.8.a). Now let p be a variable of type $\langle\rangle$ and $O!$ have type $\langle\langle\rangle\rangle$. Note that the following is a universal closure of an instance of the axiom η-Conversion (935.27) and so an axiom:

(ϑ) $\forall p(O!p \to ([\lambda\, p] = p))$

Note that φ is substitutable for p in the matrix of (ϑ) (if φ is substituted for p in $O!p \to ([\lambda\, p] = p)$, no free variable in φ gets captured) and since $\varphi\downarrow$, can instantiate (ϑ) to φ, to obtain $O!\varphi \to ([\lambda\, \varphi] = \varphi)$. Since $O!\varphi$ holds by assumption, it follows that $[\lambda\, \varphi] = \varphi$. $\bowtie$

(950.2) Let φ be any non-basic formula, i.e., any formula not in $\mathit{Base}^{\langle\rangle}$. Then by axiom (935.23), it follows that $O!\varphi$. So by (950.1), $[\lambda\, \varphi] = \varphi$. $\bowtie$

(950.3) Let φ be any formula. Then by β-Conversion (935.26), we know:

$$[\lambda\,\varphi]\!\downarrow \to ([\lambda\,\varphi] \equiv \varphi)$$

But by axiom (935.5), $[\lambda\varphi]\!\downarrow$, since $[\lambda\varphi]$ is a core λ-expression. Hence $[\lambda\varphi] \equiv \varphi$. ⋈

(950.4) It is axiomatic that $[\lambda\,\varphi]\!\downarrow$, by (935.5). So by α-Conversion (935.25), $[\lambda\,\varphi] = [\lambda\,\varphi]'$. ⋈

(950.5) We prove our theorem by the following cases: (a) φ is a constant or variable of type $\langle\,\rangle$, (b) φ is a description of type $\langle\,\rangle$, and (c) φ is a non-basic formula. However, we prove the cases in this order:

Case 1. φ is a constant or variable of type $\langle\,\rangle$

Case 2. φ is a non-basic formula.

Case 3. φ is a description of type $\langle\,\rangle$.

Case 1. φ is a constant or variable of type $\langle\,\rangle$. Then φ has no alphabetic variants and so there are no formulas of the form $\varphi\!\downarrow \to (\varphi = \varphi')$ and so nothing to prove.

Case 2. φ is a non-basic formula. Note that if we can directly show $\varphi = \varphi'$, then by propositional logic (935.1), follows that $\varphi\!\downarrow \to (\varphi = \varphi')$. Now since φ is a non-basic formula, it follows by (950.1) that $[\lambda\,\varphi] = \varphi$. But we also know by (950.4) that $[\lambda\,\varphi] = [\lambda\,\varphi]'$, where $[\lambda\,\varphi]'$ is any alphabetic variant of $[\lambda\,\varphi]$. It follows from our last two results that $\varphi = [\lambda\,\varphi]'$, by Rule =E. But by the type-theoretic definition of *alphabetic variant* (930) [16], $[\lambda\,\varphi]' = [\lambda\,\varphi']$ and so we know $\varphi = [\lambda\,\varphi']$. Note that since φ is a non-basic formula, so is φ'. So as an instance of (950.2), we have $[\lambda\,\varphi'] = \varphi'$. Hence, by Rule =E, $\varphi = \varphi'$.

Case 3. φ is a description of type $\langle\,\rangle$. Assume $\varphi\!\downarrow$. By an application of GEN to the instance of (947.1) that holds for type $\langle\,\rangle$, we know the following, where p is a variable of type $\langle\,\rangle$:

$$\forall p(p = p)$$

So by Rule ∀E, it follows that $\varphi = \varphi$. Note that if φ' is an alphabetic variant of φ, then by the definition of alphabetic variants (930) [16], $\varphi = \varphi'$ is an alphabetic variant of $\varphi = \varphi$. But $\varphi = \varphi$ is a non-basic formula and so by *Case 2*, an identity holds between it and any of its alphabetic variants. Hence, it follows that:

$$(\varphi = \varphi) = (\varphi = \varphi')$$

From this and theorem (944.2) it follows that $\Box((\varphi = \varphi) \equiv (\varphi = \varphi'))$, and so by the T schema, $(\varphi = \varphi) \equiv (\varphi = \varphi')$. And since we've established that $\varphi = \varphi$, it follows by biconditional syllogism that $\varphi = \varphi'$. ⋈

(950.6) Let φ be any formula other than a description of type $\langle\rangle$. Then by (940.1), $\varphi\downarrow$. So if φ' is any alphabetic variable of φ, it follows by (950.5) that $\varphi = \varphi'$. ⋈

(950.7) Since $\varphi\downarrow \vee \neg(\varphi\downarrow)$, we prove our theorem by cases, where the two cases are $\varphi\downarrow$ and $\neg(\varphi\downarrow)$.

Case 1. $\varphi\downarrow$. Then by (950.6), $\varphi = \varphi'$. From this and the tautology $\varphi \equiv \varphi$, it follows by Rule =E that $\varphi \equiv \varphi'$.

Case 2. $\neg(\varphi\downarrow)$. Then by the 0-ary case of axiom (935.8.a), $\neg\varphi$. Now if we can show that $\neg(\varphi')$, then by classical propositional logic, $\varphi \equiv \varphi'$. But to show $\neg(\varphi')$, it suffices, by axiom (935.8.a), to show $\neg(\varphi'\downarrow)$. So, for reductio, assume $\varphi'\downarrow$. Note that since φ' is an alphabetic variant of φ and alphabetic variance is symmetric, φ is an alphabetic variant of φ'. So by our reductio assumption and (950.5), it follows that $\varphi' = \varphi$. Hence, $\varphi\downarrow$, by (943.2), which contradicts the assumption of the present case. ⋈

(950.8) By reasoning analogous to that used in (111.6), except citing (950.3) instead of (111.2). ⋈

(951) By analogy with the proof in (114). By (950.7), we know $\vdash (\varphi \equiv \varphi')$, where φ' is any alphabetic variant of φ. Hence, by (937) [63.3], $\Gamma \vdash (\varphi \equiv \varphi')$. By definition of $\equiv$ (933.3) and Rule $\equiv_{df}$E of Definiendum Elimination (938) [90.2], it follows that $\Gamma \vdash ((\varphi \rightarrow \varphi') \,\&\, (\varphi' \rightarrow \varphi))$. So by the rules of &E (938) [86], it follows that:

(A) $\Gamma \vdash (\varphi \rightarrow \varphi')$

(B) $\Gamma \vdash (\varphi' \rightarrow \varphi)$

Now to justify the left-to-right direction of the Rule of Alphabetic Variants, assume $\Gamma \vdash \varphi$. Then from this and (A), it follows by (937) [63.5] that $\Gamma \vdash \varphi'$. By analogous reasoning from (B), if $\Gamma \vdash \varphi'$, then $\Gamma \vdash \varphi$. ⋈

(952.1) – **(952.4)** (Exercises)

(953.1) – **(953.2)** (Exercises)

(954.1) – **(954.3)** (Exercises)

(955.1) – **(955.4)**, **(955.7)** – **(955.8)** (Exercises)

(956.1) – **(956.5)** (Exercises)

(957) (Exercises)

(958.1) For reductio, assume $\exists F(F\imath p(p \,\&\, \neg p))$, and suppose P is such a property, so that we know $P\imath p(p \,\&\, \neg p)$, where P has type $\langle\langle\rangle\rangle$ and p has type $\langle\rangle$. Note that the modally strict version of Russell's analysis, which was derived in type-theoretic form as theorem (957) [151], asserts the following for any type t and variables x and z of type t:

$\psi_x^{\imath x\varphi} \equiv \exists x(\mathcal{A}\varphi \ \&\ \forall z(\mathcal{A}\varphi_x^z \to z = x) \ \&\ \psi)$, provided (a) ψ is either an exemplification formula $\Pi^n\kappa_1 \ldots \kappa_n$ $(n \geq 1)$ or an encoding formula $\kappa_1 \ldots \kappa_n\Pi^n$ $(n \geq 1)$, (b) x occurs in ψ and only as one or more of the κ_i $(1 \leq i \leq n)$, and (c) z is substitutable for x in φ and doesn't occur free in φ

So if we let t be the type $\langle\,\rangle$, x be a variable of this type, and ψ be the formula Px, then the above theorem has the following instance:

$$P\imath x(p \ \&\ \neg p) \equiv \exists x(\mathcal{A}(p \ \&\ \neg p) \ \&\ \forall z(\mathcal{A}(p \ \&\ \neg p) \to z = x) \ \&\ Px)$$

So it follows that $\exists x(\mathcal{A}(p \ \&\ \neg p) \ \&\ \forall z(\mathcal{A}(p \ \&\ \neg p) \to z = x) \ \&\ Px)$. Let a be such an object, so that we know:

$$\mathcal{A}(p \ \&\ \neg p) \ \&\ \forall z(\mathcal{A}(p \ \&\ \neg p) \to z = a) \ \&\ Pa$$

Then $\mathcal{A}(p \ \&\ \neg p)$. But $\neg(p \ \&\ \neg p)$ is a theorem, and so by Rule RA (956.2) [135], $\mathcal{A}\neg(p \ \&\ \neg p)$. But then by axiom (935.11), $\neg\mathcal{A}(p \ \&\ \neg p)$. Contradiction. ⋈

(958.2) By definition of $\downarrow$ for propositions (933.6.b), we know $\imath p(p \ \&\ \neg p)\downarrow \equiv \exists F(F\imath p(p \ \&\ \neg p))$. But from this and the previous theorem (958.1), it follows that $\neg(\imath p(p \ \&\ \neg p)\downarrow)$. ⋈

(958.3) Axiom (935.8.a) asserts $\Pi \to \Pi\downarrow$, where Π is any term of type $\langle\,\rangle$. So, where p is a variable of type $\langle\,\rangle$, the following is an instance: $\imath p(p \ \&\ \neg p) \to \imath p(p \ \&\ \neg p)\downarrow$. (This asserts: if the proposition that is both true and not true is true, then the proposition that is both true and not true exists.) But the previous theorem is $\neg(\imath p(p \ \&\ \neg p)\downarrow)$. Hence $\neg\imath p(p \ \&\ \neg p)$. ⋈

(958.4) Let p be a variable of type $\langle\,\rangle$. Since p is a formula, let ψ be the formula p and consider the formula $\psi_p^{\imath p\varphi}$, where φ is any formula. As an instance of the type-theoretic, modally-strict version of Russell's analysis of descriptions (957) [151], we know:

$$\psi_p^{\imath p\varphi} \equiv \exists p(\mathcal{A}\varphi \ \&\ \forall q(\mathcal{A}\varphi_p^q \to q = p) \ \&\ p)$$

But $\psi_p^{\imath p\varphi}$ is simply the formula $\imath p\varphi$. Hence:

$$\imath p\varphi \equiv \exists p(\mathcal{A}\varphi \ \&\ \forall q(\mathcal{A}\varphi_p^q \to q = p) \ \&\ p) \qquad ⋈$$

(959.1) – **(959.7)** (Exercises)

(960.1) (Exercises)

(960.3) – **(960.4)** (Exercises)

(961.1) The reasoning is a variant of that used in (189) and (245.1), but typed. Let t be any type, x be a variable of type t, and F and G be variables of type $\langle t\rangle$. Assume $\forall x(xF \equiv xG)$. By BF, it suffices to show $\forall x\Box(xF \equiv xG)$, and by GEN, it suffices to show $\Box(xF \equiv xG)$. But we know the following:

(ϑ) $xF \equiv \Box xF$ instance of (959.7) [179.2]

(ζ) $xG \equiv \Box xG$ instance of (959.7) [179.2]

(ξ) $\Box(xF \equiv xG) \equiv (\Box xF \equiv \Box xG)$ instance of (959.7) [179.5]

where x, F, and G have the types indicated above. So if we can establish the right side of (ξ), we're done. We do this as follows:

$$\begin{array}{lcll} \Box xF & \equiv & xF & \text{by } (\vartheta) \\ & \equiv & xG & \text{by assumption} \\ & \equiv & \Box xG & \text{by } (\zeta) \qquad \bowtie \end{array}$$

(961.2) Assume $O!F$ and $O!G$. Further assume $\forall x(xF \equiv xG)$. Then by (961.1), the latter implies $\Box\forall x(xF \equiv xG)$. So by definition (933.10), $F = G$. $\bowtie$

(961.3) (Exercise)

(961.4) Assume $A!F$, $A!G$, and $\forall\mathcal{H}(F\mathcal{H} \equiv G\mathcal{H})$. Then by (961.3), the latter implies $\Box\forall\mathcal{H}(F\mathcal{H} \equiv G\mathcal{H})$. But our first two assumptions and this last result imply, by definition (933.10), $F = G$. $\bowtie$

(961.5) Let $t_1, \ldots, t_n$ be any types ($n \geq 1$), and let:

(A) F be a variable of type $\langle t_1, \ldots, t_n\rangle$,

(B) $O!$ be a defined property having type $\langle\langle t_1, \ldots, t_n\rangle\rangle$,

(C) $x_1, \ldots, x_n$ be variables having types $t_1, \ldots, t_n$, respectively, and

(D) φ be any formula such that (a) F doesn't occur free in φ and (b) $x_1, \ldots, x_n$ don't occur in encoding position in φ [930] (9.1).

By (C) and (D.b), it follows that $[\lambda x_1 \ldots x_n\, \varphi]$ is a core λ-expression. Hence $[\lambda x_1 \ldots x_n\, \varphi]\downarrow$. From this it follows, by (935.22) and (B), that:

(ϑ) $O![\lambda x_1 \ldots x_n\, \varphi]$

and, by β-Conversion (935.26), that:

$$[\lambda x_1 \ldots x_n\, \varphi]x_1 \ldots x_n \equiv \varphi$$

From the latter it follows by n applications of GEN:

$$\forall x_1 \ldots \forall x_n([\lambda x_1 \ldots x_n\, \varphi]x_1 \ldots x_n \equiv \varphi)$$

Since this claim has been established by a modally strict proof, it follows by RN (68) that:

$$\Box\forall x_1 \ldots \forall x_n([\lambda x_1 \ldots x_n\, \varphi]x_1 \ldots x_n \equiv \varphi)$$

Conjoining (ϑ) with this last result, we have:

(ξ) $O![\lambda x_1 \ldots x_n\, \varphi] \mathbin{\&} \Box\forall x_1 \ldots \forall x_n([\lambda x_1 \ldots x_n\, \varphi]x_1 \ldots x_n \equiv \varphi)$

By (D.a) above, F doesn't occur free in φ, and we can use Rule ∃I (939) [101.1] to existentially generalize and conclude:

$$\exists F^n(O!F \mathbin{\&} \Box\forall x_1 \ldots \forall x_n(F^n x_1 \ldots x_n \equiv \varphi)) \qquad \bowtie$$

(961.7) Let p be a variable of type $\langle\rangle$, $O!$ be the defined relation term having type $\langle\langle\rangle\rangle$, and let φ be any formula in which p doesn't occur free. Note that it is axiomatic that $O![\lambda\, \varphi]$ (935.23). Independently, by applying RN to theorem (950.3), we know: $\Box([\lambda\, \varphi] \equiv \varphi)$. Thus, we know:

$$O![\lambda\, \varphi] \mathbin{\&} \Box([\lambda\, \varphi] \equiv \varphi)$$

But by (935.5), we know $[\lambda\, \varphi]{\downarrow}$ (this also follows from the previously established fact that $O![\lambda\, \varphi]$). Since p doesn't occur free in φ, we may use Rule ∃I (939) [101.1] and conclude $\exists p(O!p \mathbin{\&} \Box(p \equiv \varphi))$. $\bowtie$

(961.8) (Exercises)

(962.1) We use reasoning analogous to that used in (189), (245.1), but typed as in (961.1). Let t be any type, x and y be variables of type t, and F be a variable of type $\langle t\rangle$. Assume $\forall F(xF \equiv yF)$. By BF, it suffices to show $\forall x\Box(xF \equiv yF)$, and by GEN, it suffices to show $\Box(xF \equiv yF)$. Now we know:

(ϑ)	$xF \equiv \Box xF$	instance of (959.7) [179.2]
(ζ)	$yF \equiv \Box yF$	instance of (959.7) [179.2]
(ξ)	$\Box(xF \equiv yF) \equiv (\Box xF \equiv \Box yF)$	instance of (959.7) [179.5]

where x, y, and F have the types indicated above. So if we can establish the right side of (ξ), we're done. We do this as follows:

$$\begin{array}{lcll} \Box xF & \equiv & xF & \text{by } (\vartheta) \\ & \equiv & yF & \text{by assumption} \\ & \equiv & \Box yF & \text{by } (\zeta) \qquad \bowtie \end{array}$$

(962.2) Let x and y be variables of the same type. Then we have four cases:

(A) x and y are variables of type i

(B) x and y are variables having a type of the form $\langle t\rangle$, for some type t

(C) x and y are variables having a type of the form $\langle t_1,\ldots,t_n\rangle$, for some types $t_1,\ldots,t_n$ ($n \geq 2$)

(D) x and y are variables of type $\langle\,\rangle$

Case (B) has already been established as theorem (961.4). So it remains to show cases (A), (C), and (D).

(A) Assume $A!x$, $A!y$, and $\forall F(xF \equiv yF)$, where $A!$ and F have type $\langle i \rangle$. By the relevant instance of (962.1), the third assumption implies $\Box\forall F(xF \equiv yF)$. But our first two assumptions and this last result imply, by definition (933.9), $x = y$.

(C) Assume $A!x$, $A!y$, and $\forall F(xF \equiv yF)$, where $A!$ and F have type $\langle\langle t_1, \ldots, t_n\rangle\rangle$. By the relevant instance of (962.1), the third assumption implies $\Box\forall F(xF \equiv yF)$. But our first two assumptions and this last result imply, by definition (933.11), $x = y$.

(D) Assume $A!x$, $A!y$, and $\forall F(xF \equiv yF)$, where $A!$ and F have type $\langle\langle\,\rangle\rangle$. By the relevant instance of (962.1), the third assumption implies $\Box\forall F(xF \equiv yF)$. But our first two assumptions and this last result imply, by definition (933.12), $x = y$. ⋈

(962.3) – **(962.8)** (Exercises)

(963.1) Let t be any type, x be a variable of type t, $A!$ have type $\langle t \rangle$, F be variable having type $\langle t \rangle$, and φ be any formula in which x doesn't occur free. Then by (935.32):

$$\exists x(A!x \,\&\, \forall F(xF \equiv \varphi))$$

Let a be a arbitrary such object of type t, so that we know:

(ϑ) $A!a \,\&\, \forall F(aF \equiv \varphi)$

Then to show that a is a witness to our theorem, it remains, by the definition of the uniqueness quantifier and Rule EI, only to show that:

$$\forall x(A!x \,\&\, \forall F(xF \equiv \varphi) \rightarrow x = a)$$

So by GEN, assume both $A!x$ and $\forall F(xF \equiv \varphi)$, to show $x = a$. Given the former, we now know both a and x are abstract, and from the latter and (ϑ), it follows that:

$$\forall F(xF \equiv aF)$$

Now note that the proof of (962.2) goes exactly like the proof of the second-order version (245.2). As part of the proof of (245.2), we established that (245.1), i.e., that $\forall F(xF \equiv yF)$ implies $\Box\forall F(xF \equiv yF)$. This holds matter what type x and y are; the reasoning in (245.1) converts to a proof of (962.1), axiom (935.30) and the theorems in (959.7) [179]. So, no matter what type x and a have, it follows from $A!x$, $A!a$, and $\forall F(xF \equiv aF)$ that:

(ζ) $\Box\forall F(xF \equiv aF)$

It remains to show that (ζ) implies $x = a$ for every type t. But t falls into one of the following mutually exclusive and jointly exhaustive cases:

- $t = i$
- $t = \langle t' \rangle$, for some type t'
- $t = \langle t'_1, \ldots, t'_n \rangle$, for some types $t'_1, \ldots, t'_n$, where $n \geq 2$
- $t = \langle\,\rangle$

Note that these four cases correspond to the four cases of the definition of identity for abstract objects, (933.9) – (933.12). In each of these cases, the definitions imply that abstract objects x and y, of any type t, which necessarily encode the same properties are identical. Hence, in each case, it follows from (ζ) that $x = a$. ⋈

(963.3) Let t be any type, α be a variable of type t, $A!$ have type $\langle t \rangle$, F be variable having type $\langle t \rangle$, and φ be any formula in which α doesn't occur free. Then (963.1) holds and by the Rule of Actualization:

$$\mathcal{A}\exists!\alpha(A!\alpha \,\&\, \forall F(\alpha F \equiv \varphi))$$

So by (959.7) [176.2]:

$$\imath\alpha(A!\alpha \,\&\, \forall F(\alpha F \equiv \varphi))\downarrow \qquad ⋈$$

(963.4) (Exercise)

(963.5) The proofs of the typed versions of the theorems listed carry over from second-order object theory.

(964.1) (Exercise)

(964.2) Suppose, for reductio, that $\imath p\varphi\downarrow \to \imath p\varphi$. Then let ψ be any formula in which p doesn't occur free and consider the canonical description:

$$\imath p(A!p \,\&\, \forall F(pF \equiv \psi))$$

Let's abbreviate this description as $\imath p\chi$. By (963.3):

(ϑ) $\imath p\chi\downarrow$

It therefore follows from our reductio hypothesis that $\imath p\chi$. But when t is the type $\langle\,\rangle$, $A!p \to \neg p$ is an instance of axiom (935.24), and since $\forall p(A!p \to \neg p)$ is a closure, it is also an axiom. Since (ϑ) tells us that $\imath p\chi$ is significant, we may instantiate it into $\forall p(A!p \to \neg p)$ to obtain $A!\imath p\chi \to \neg\imath p\chi$. But since $\imath p\chi\downarrow$, the antecedent is derivable (a) by setting φ to ψ in (963.4), (b) universally generalizing the result, (c) instantiating to $\imath p\chi$ to obtain $\imath p\chi = \imath p\chi \to A!\imath p\chi$, and then (d) deriving $A!\imath p\chi$ from the fact that $\imath p\chi = \imath p\chi$. Hence $\neg\imath p\chi$. Contradiction. ⋈

(965.1) The typed versions of the proofs of (268.1) – (268.3) remain proofs in typed object theory. ⋈

(965.2) By a proof analogous to (269). ⋈

(965.3) – **(965.4)** By proofs analogous to (271.2) and (272.2), respectively. ⋈

(965.5) Let w, x, y be variables of type t, z be a variable of type t', and F be a variable of type $\langle t \rangle$. Assume $[\lambda z\, \varphi]\downarrow$, where no free occurrence of x in φ is in encoding position [930] (9.1). Note that the condition on x in φ implies that the term $[\lambda x\, w[\lambda z\, \varphi]]$ is a core λ-expression [930] (9.2). So $[\lambda x\, w[\lambda z\, \varphi]]\downarrow$ is an instance of axiom (935.5). Now consider the unary instance of the Corollary to Kirchner Theorem (965.4):

$$[\lambda x\, w[\lambda z\, \varphi]]\downarrow \to \forall x \forall y(\forall F(Fx \equiv Fy) \to \Box(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi^y_x]))$$

Since the antecedent is axiomatic, we may infer:

$$\forall x \forall y(\forall F(Fx \equiv Fy) \to \Box(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi^y_x]))$$

So for an arbitrary x and y, it follows that:

$$\forall F(Fx \equiv Fy) \to \Box(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi^y_x])$$

Since we derived the above from no assumptions with w free, it follows by GEN:

$$\forall w(\forall F(Fx \equiv Fy) \to \Box(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi^y_x]))$$

So by (95.2):

$$\forall F(Fx \equiv Fy) \to \forall w \Box(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi^y_x])$$

By a Rule of Substitution and the modally strict equivalence of the Barcan Formulas (167.1) and (167.2):

$$(\vartheta)\quad \forall F(Fx \equiv Fy) \to \Box \forall w(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi^y_x])$$

Now our goal is to show that we can derive the following from our initial assumption:

$$\forall F(Fx \equiv Fy) \to [\lambda z\, \varphi] = [\lambda z\, \varphi^y_x]$$

So assume $\forall F(Fx \equiv Fy)$ (as a local assumption). It then follows from (ϑ) that $\Box \forall w(w[\lambda z\, \varphi] \equiv w[\lambda z\, \varphi^y_x])$. Now let $O!$ be a defined constant of type $\langle\langle t' \rangle\rangle$. Then it is straightforward to derive, from definition (933.10):

(ξ) $(O!{\scriptstyle[\lambda z\,\varphi]}\ \&\ O!{\scriptstyle[\lambda z\,\varphi^y_x]}) \to ([\lambda z\,\varphi] = [\lambda z\,\varphi^y_x] \equiv \Box\forall w(w[\lambda z\,\varphi] \equiv w[\lambda z\,\varphi^y_x]))$

But the first conjunct of the antecedent follows from our initial assumption, for $[\lambda z\,\varphi]\!\downarrow\, \to O!{\scriptstyle[\lambda z\,\varphi]}$ is an instance of axiom (935.22). And we can derive the second conjunct of the antecedent from what we know:

> By our initial assumption, $[\lambda z\,\varphi]\!\downarrow$. But given the condition on x in φ, we also know that $[\lambda x\,[\lambda z\,\varphi]\!\downarrow]\!\downarrow$. So we can instantiate $[\lambda x\,[\lambda z\,\varphi]\!\downarrow]$ into our local assumption $\forall F(Fx \equiv Fy)$, to obtain $[\lambda x\,[\lambda z\,\varphi]\!\downarrow]x \equiv [\lambda x\,[\lambda z\,\varphi]\!\downarrow]y$. Applying β-Conversion to both sides, this reduces to $[\lambda z\,\varphi]\!\downarrow\, \equiv [\lambda z\,\varphi]\!\downarrow^y_x$. Hence, $[\lambda z\,\varphi]\!\downarrow^y_x$, i.e., $[\lambda z\,\varphi^y_x]\!\downarrow$. So by the relevant instance of axiom (935.22), $O!{\scriptstyle[\lambda z\,\varphi^y_x]}$.

Since we've established both conjuncts of the antecedent of (ξ), it follows that:

$$[\lambda z\,\varphi] = [\lambda z\,\varphi^y_x] \equiv \Box\forall w(w[\lambda z\,\varphi] \equiv w[\lambda z\,\varphi^y_x])$$

And since we've established the right-hand side, it follows that $[\lambda z\,\varphi] = [\lambda z\,\varphi^y_x]$.
$\bowtie$

(965.6) Let t be any type, x be a variable of type t and F a variable of type $\langle t\rangle$. Assume $\forall x([\lambda F\, xF]\!\downarrow)$, for reductio. Now consider the type $\langle t\rangle$. Since we can form an instance of theorem (965.2) in which the x and y are variables of type $\langle t\rangle$, we can express such an instance by using F and G as variables of $\langle t\rangle$, $A!$ as a constant of type $\langle\langle t\rangle\rangle$ and $\mathcal{H}$ as a variable of type $\langle\langle t\rangle\rangle$, thereby obtaining the following:

$$\exists F\exists G(A!{\scriptstyle F}\ \&\ A!{\scriptstyle G}\ \&\ F \neq G\ \&\ \forall\mathcal{H}(\mathcal{H}{\scriptstyle F} \equiv \mathcal{H}{\scriptstyle G}))$$

Let P and Q be such properties, so that we know:

$$A!{\scriptstyle P}\ \&\ A!{\scriptstyle Q}\ \&\ P \neq Q\ \&\ \forall\mathcal{H}(\mathcal{H}{\scriptstyle P} \equiv \mathcal{H}{\scriptstyle Q})$$

So, by comprehension for abstract objects of type t, there is an abstract object of type t that encodes just P. Suppose a is such an object, so that we know:

(ϑ) $aP\ \&\ \neg aQ$

Then by our reductio assumption, $[\lambda F\, aF]\!\downarrow$. So where φ is aF, we have the following unary instance of the Corollary to the Kirchner Theorem (965.4):

$$[\lambda F\, aF]\!\downarrow\, \to \forall F\forall G(\forall\mathcal{H}(\mathcal{H}{\scriptstyle F} \equiv \mathcal{H}{\scriptstyle G}) \to \Box(aF \equiv aG))$$

Since we know the antecedent, it follows that:

$$\forall F\forall G(\forall\mathcal{H}(\mathcal{H}{\scriptstyle F} \equiv \mathcal{H}{\scriptstyle G}) \to \Box(aF \equiv aG))$$

By instantiating the above to P and Q, we obtain:

$$\forall\mathcal{H}(\mathcal{H}P \equiv \mathcal{H}Q) \rightarrow \Box(aP \equiv aQ))$$

But we know the antecedent, and so:

$$\Box(aP \equiv aQ)$$

Hence $aP \equiv aQ$, which contradicts (ϑ). ⋈

(**966.1**) Let x, F, and $O!$ have the types indicated by the statement of the theorem. By the **Observation** in (930), we know that $[\lambda F\, O!F \,\&\, x[\lambda z\, Fz]]$ is a core λ-expression. So as an instance of axiom (935.5), we know:

$$[\lambda F\, O!F \,\&\, x[\lambda z\, Fz]]{\downarrow}$$

We can't *directly* infer from this that $[\lambda F\, O!F \,\&\, xF]{\downarrow}$ since η-Conversion yields that $[\lambda z\, Fz] = F$ only under the condition that $O!F$. But we can establish that $[\lambda F\, O!F \,\&\, xF]{\downarrow}$ by appealing to the following instance of axiom (935.28):

$$([\lambda F\, O!F \,\&\, x[\lambda z\, Fz]]{\downarrow} \,\&\, \Box\forall F((O!F \,\&\, x[\lambda z\, Fz]) \equiv (O!F \,\&\, xF))) \rightarrow$$
$$[\lambda F\, O!F \,\&\, xF]{\downarrow}$$

So to prove our theorem, it remains only to show:

$$\Box\forall F((O!F \,\&\, x[\lambda z\, Fz]) \equiv (O!F \,\&\, xF))$$

By GEN and RN, it suffices to show:

$$(O!F \,\&\, x[\lambda z\, Fz]) \equiv (O!F \,\&\, xF)$$

So by propositional logic, we have to show:

$$O!F \rightarrow (x[\lambda z\, Fz] \equiv xF)$$

So assume $O!F$. Then by η-Conversion (935.27), $[\lambda z\, Fz] = F$. But then from the tautology $x[\lambda z\, Fz] \equiv x[\lambda z\, Fz]$, it follows that $x[\lambda z\, Fz] \equiv xF$. ⋈

(**966.2**) Let x be a variable of type t, F and G be variables of type $\langle t\rangle$, $O!$ be a defined constant of type $\langle\langle t\rangle\rangle$, and $\mathcal{H}$ be a variable of type $\langle\langle t\rangle\rangle$. Now take $O!F$ and $O!G$ as global assumptions.

($\rightarrow$) As a local assumption, suppose:

(ϑ) $\Box\forall x(xF \equiv xG)$

It then follows from our global assumptions and local assumption that $F = G$, by definition (933.10). Hence from the logical theorem that $\Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}F)$, it follows that $\Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G)$, by Rule =E.[484]

($\leftarrow$) Now let $\Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G)$ be our local assumption. To show $\Box\forall x(xF \equiv xG)$, first note that the following is an instance of Kirchner's Theorem (965.2), raised to the type $\langle t\rangle$:

[484] There is also a proof of this direction that doesn't use either definition of of $F = G$ (933.10) or the fact that F and G are ordinary. That is, (ϑ) is sufficient to derive $\Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G)$. Consider the

$$[\lambda F\, O!F \,\&\, xF]\downarrow \equiv \Box\forall F\forall G(\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G) \rightarrow ((O!F \,\&\, xF) \equiv (O!G \,\&\, xG)))$$

By (966.1), the left condition is a theorem, and so:

(ζ) $\Box\forall F\forall G(\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G) \rightarrow ((O!F \,\&\, xF) \equiv (O!G \,\&\, xG)))$

By now familiar modal reasoning using the Converse Barcan Formula, it follows that:

$$\forall F\forall G\Box(\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G) \rightarrow ((O!F \,\&\, xF) \equiv (O!G \,\&\, xG)))$$

Hence by $\forall$E:

$$\Box(\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G) \rightarrow ((O!F \,\&\, xF) \equiv (O!G \,\&\, xG)))$$

By the K axiom, this implies:

$$\Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G) \rightarrow \Box((O!F \,\&\, xF) \equiv (O!G \,\&\, xG))$$

From this and our local assumption, it follows that:

$$\Box((O!F \,\&\, xF) \equiv (O!G \,\&\, xG))$$

Now the derivation $\Box((\varphi \,\&\, \psi) \equiv (\chi \,\&\, \theta)) \vdash \Box((\varphi \,\&\, \chi) \rightarrow (\psi \equiv \theta))$ is valid:

Proof. We leave it as an exercise to show that there is a modally strict derivation of $(\varphi \,\&\, \chi) \rightarrow (\psi \equiv \theta)$ from the premise $(\varphi \,\&\, \psi) \equiv (\chi \,\&\, \theta)$. Hence by RN, there is a derivation of necessitation of the conclusion from the necessitation of the premise.

So from our last result, it follows that:

$$\Box((O!F \,\&\, O!G) \rightarrow (xF \equiv xG))$$

So by the K axiom:

(ξ) $\Box(O!F \,\&\, O!G) \rightarrow \Box(xF \equiv xG)$

But both $O!F$ and $O!G$ are global assumptions. And it is easy to show both:

---

following instance of Comprehension for Abstract Objects:

$$\exists x(A!x \,\&\, \forall K(xK \equiv \Box\forall\mathcal{H}(\mathcal{H}K \equiv \mathcal{H}G)))$$

Let a be such an object, so that we know:

(ξ) $A!a \,\&\, \forall K(aK \equiv \Box\forall\mathcal{H}(\mathcal{H}K \equiv \mathcal{H}G))$

If we instantiate the second conjunct to G, we therefore know:

$$aG \equiv \Box\forall\mathcal{H}(\mathcal{H}G \equiv \mathcal{H}G)$$

But it easy to establish the right side as a theorem – it follows by GEN and RN from the tautology $\mathcal{H}G \equiv \mathcal{H}G$. Hence, aG. But (ϑ) implies, by the T schema, $\forall x(xF \equiv xG)$, and so implies $aF \equiv aG$, by $\forall$E. Hence aF, by biconditional syllogism. But if we instantiate the second conjunct of (ξ) to F, we have $aF \equiv \Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G)$. Hence $\Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G)$.

$O!F \equiv \Box O!F$

$O!G \equiv \Box O!G$

Proof. The left-to-right direction of each of the above follows by (959.7) [180.1], and the right-to-left direction is an intance of the T schema.

Hence, we know both $\Box O!F$ and $\Box O!G$, from which we may infer $\Box(O!F \,\&\, O!G)$ by familiar modal reasoning. Then by (ξ), $\Box(xF \equiv xG)$. Since x isn't free in any assumption, we may infer by GEN:

$\forall x \Box(xF \equiv xG)$

So by the Barcan formula, $\Box\forall x(xF \equiv xG)$. ⋈

(966.3) Let t be any type, F and G be variables of type $\langle t\rangle$, $O!$ be a defined constant of type $\langle\langle t\rangle\rangle$, and $\mathcal{H}$ be a variable of type $\langle\langle t\rangle\rangle$. Then assume $O!F$, $O!G$), and $\Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G)$. Then where x is a variable of type t, it follows By (966.2) that $\Box\forall x(xF \equiv xG)$. Hence, we've established the left disjunct of the definiens of the definition of $F = G$ (933.10). ⋈

(967.2) Let t be any type, and F and G be variables of type $\langle t\rangle$, $O!$ be a defined constant of type $\langle\langle t\rangle\rangle$, and $\mathcal{H}$ a variable of type $\langle\langle t\rangle\rangle$. Then we derive our theorem from an instance of the safe extension axiom (935.28), by establishing the antecedent of that instance. The following is an instance of axiom (935.28):

$$
\begin{aligned}
&([\lambda FG\, O!F \,\&\, O!G \,\&\, \Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G)]{\downarrow} \;\&\\
&\quad \Box\forall F\forall G((O!F \,\&\, O!G \,\&\, \Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G)) \equiv (O!F \,\&\, O!G \,\&\, \Box\forall x(xF \equiv xG)))) \to\\
&\quad [\lambda FG\, O!F \,\&\, O!G \,\&\, \forall x(xF \equiv xG)]{\downarrow}
\end{aligned}
$$

So to prove our theorem, it suffices to show both conjuncts of the antecedent. But in the first conjunct, the term:

$[\lambda FG\, O!F \,\&\, O!G \,\&\, \Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G)]$

is a core λ-expression. So the first conjunct is axiomatic. It therefore remains only to show the second conjunct, i.e.,

$\Box\forall F\forall G((O!F \,\&\, O!G \,\&\, \Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G)) \equiv (O!F \,\&\, O!G \,\&\, \Box\forall x(xF \equiv xG)))$

By GEN and RN, it suffices to show:

$(O!F \,\&\, O!G \,\&\, \Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G)) \equiv (O!F \,\&\, O!G \,\&\, \Box\forall x(xF \equiv xG))$

So by propositional logic, we have to show:

$(O!F \,\&\, O!G) \to (\Box\forall\mathcal{H}(\mathcal{H}F \equiv \mathcal{H}G) \equiv \Box\forall x(xF \equiv xG))$

But this is an immediate consequence of (966.2). ⋈

(967.2) [Alternative Proof:] We derive our theorem from an instance of the safe extension axiom (935.28), by establishing the antecedent of that instance. The following is an instance of axiom (935.28):

$$([\lambda FG\, O!F \,\&\, O!G \,\&\, \Box\forall x(x[\lambda z\, Fz] \equiv x[\lambda z\, Gz])]\downarrow \,\& \\ \Box\forall F\forall G((O!F \,\&\, O!G \,\&\, \Box\forall x(x[\lambda z\, Fz] \equiv x[\lambda z\, Gz])) \equiv (O!F \,\&\, O!G \,\&\, \Box\forall x(xF \equiv xG)))) \to \\ [\lambda FG\, O!F \,\&\, O!G \,\&\, \forall x(xF \equiv xG)]\downarrow$$

So to prove our theorem, it suffices to show both conjuncts of the antecedent. But in the first conjunct, the term:

$$[\lambda FG\, O!F \,\&\, O!G \,\&\, \Box\forall x(x[\lambda z\, Fz] \equiv x[\lambda z\, Gz])]$$

is a core λ-expression. So the first conjunct is axiomatic. It therefore remains only to show the second conjunct, i.e.,

$$\Box\forall F\forall G((O!F \,\&\, O!G \,\&\, \Box\forall x(x[\lambda z\, Fz] \equiv x[\lambda z\, Gz])) \equiv (O!F \,\&\, O!G \,\&\, \Box\forall x(xF \equiv xG)))$$

By GEN and RN, it suffices to show:

$$(O!F \,\&\, O!G \,\&\, \Box\forall x(x[\lambda z\, Fz] \equiv x[\lambda z\, Gz])) \equiv (O!F \,\&\, O!G \,\&\, \Box\forall x(xF \equiv xG))$$

So by propositional logic, we have to show:

$$(O!F \,\&\, O!G) \to (\Box\forall x(x[\lambda z\, Fz] \equiv x[\lambda z\, Gz]) \equiv \Box\forall x(xF \equiv xG))$$

So assume $O!F$ and $O!G$. Then by η-Conversion (935.27), we know both $[\lambda zFz] = F$ and $[\lambda z\, Gz] = G$. ($\to$) Assume $\Box\forall x(x[\lambda z\, Fz] \equiv x[\lambda z\, Gz])$. Then by two applications of Rule =E, $\Box\forall x(xF \equiv xG)$. ($\leftarrow$) Assume $\Box\forall x(xF \equiv xG)$. Then by the symmetry of identity and two applications of Rule =E, $\Box\forall x(x[\lambda z\, Fz] \equiv x[\lambda z\, Gz])$. ⋈

(967.4) (Exercise)

(967.6) (Exercise)

(967.8) (Exercise)

(968.1) We prove our theorem by cases:

- $O!x \to x =_E x$, where x has type i and $O!$ has type $\langle i\rangle$
- $O!F \to F =_E F$, where F has type $\langle t\rangle$, for any type t, and $O!$ has type $\langle\langle t\rangle\rangle$
- $O!F \to F =_E F$, where F has type $\langle t_1, \ldots, t_n\rangle$ ($n \geq 2$), for any types $t_1, \ldots, t_n$, and $O!$ has type $\langle\langle t_1, \ldots, t_n\rangle\rangle$
- $O!p \to p =_E p$, where p has type $\langle\,\rangle$ and $O!$ has type $\langle\langle\,\rangle\rangle$.

Case 1. Show $O!x \to x =_E x$, where x has type i and $O!$ has type $\langle i \rangle$. Assume $O!x$. Then by the idempotence of & (938) [85.6], $O!x \,\&\, O!x$. Now where F is a variable type $\langle t \rangle$, it is a tautology that $Fx \equiv Fx$. Hence by GEN and RN, $\Box\forall x(Fx \equiv Fx)$. So by &I (938) [86.1], $O!x \,\&\, O!x \,\&\, \Box\forall x(Fx \equiv Fx)$. Independently, $[\lambda xy\, O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)]$ is a core λ-expression, we know by axiom (935.5) that:

$$[\lambda xy\, O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)]\downarrow$$

So, independently, by β-Conversion (935.26), it is a theorem that:

$$[\lambda xy\, O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)]xy \equiv (O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy))$$

So by GEN:

$$\forall x\forall y([\lambda xy\, O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)]xy \equiv (O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)))$$

Instantiating both universal quantifiers to x:

$$[\lambda xy\, O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)]xx \equiv (O!x \,\&\, O!x \,\&\, \Box\forall F(Fx \equiv Fx))$$

Since we've established the right side, it follows that:

$$[\lambda xy\, O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)]xx$$

Hence by definition (967.1) and infix notation, $x =_E x$. ⋈

Case 2. Show $O!\mathit{F} \to F =_E F$, where F has type t, for any type t, and $O!$ has type $\langle t \rangle$. Assume $O!\mathit{F}$. Then by the idempotence of & (938) [85.6], $O!\mathit{F} \,\&\, O!\mathit{F}$. Now where x is a variable type t', it is a tautology that $xF \equiv xF$. Hence by GEN and RN, $\Box\forall x(xF \equiv xF)$. So $O!\mathit{F} \,\&\, O!\mathit{F} \,\&\, \Box\forall x(xF \equiv xF)$. Independently, by (967.2), it is a theorem that:

$$[\lambda FG\, O!F \,\&\, O!G \,\&\, \Box\forall x(xF \equiv xG)]\downarrow$$

We leave it as an exercise to show, by an appeal to β-Conversion (935.26), that it is a theorem that:

$$[\lambda FG\, O!F \,\&\, O!G \,\&\, \Box\forall x(xF \equiv xG)]FF \equiv (O!F \,\&\, O!F \,\&\, \Box\forall x(xF \equiv xF))$$

Since we've established the right side, it follows that:

$$[\lambda FG\, O!F \,\&\, O!G \,\&\, \Box\forall x(xF \equiv xG)]FF$$

Hence by definition (967.3) and infix notation, $F =_E F$. ⋈

Case 3. (Exercise)

Case 4. Show $O!p \to p =_E p$, where p has type $\langle\,\rangle$ and $O!$ has type $\langle\langle\,\rangle\rangle$. So assume $O!p$, where $O!$ has type $\langle\langle\,\rangle\rangle$. Then by the idempotence of & (938) [85.6], $O!p \,\&\, O!p$. But now consider the property $[\lambda x\, p]$, where x is a variable of type i. By axiom (935.22), we know $O![\lambda x\, p]$, where $O!$ has type $\langle\langle i \rangle\rangle$. Then by Case 2, we've established $[\lambda x\, p] =_E [\lambda x\, p]$. Hence, $O!p \,\&\, O!p \,\&\, [\lambda x\, p] =_E [\lambda x\, p]$. By now familiar reasoning, we independently know:

$$[\lambda pq\, O!p \;\&\; O!q \;\&\; [\lambda x\, p] =_E [\lambda x\, q]]\downarrow$$

Hence, by β-Conversion:

$$[\lambda pq\, O!p \;\&\; O!q \;\&\; [\lambda x\, p] =_E [\lambda x\, q]]pp \equiv (O!p \;\&\; O!p \;\&\; [\lambda x\, p] =_E [\lambda x\, p])$$

Since we've established the right side:

$$[\lambda pq\, O!p \;\&\; O!q \;\&\; [\lambda x\, p] =_E [\lambda x\, q]]pp$$

Hence by definition (967.7) and infix notation, $p =_E p$ ⋈

(968.3) – **(968.5)** (Exercises)

(971.5) Follow the proof in (486.1), but first (a) formulate a type-theoretic version of the notion of a condition on propositional properties, as defined in (480), (b) prove a type-theoretic version of theorem (482.1), which asserts $\exists x(\mathit{Situation}(x) \& \forall F(xF \equiv \varphi))$, provided φ is a condition on propositional properties in which x doesn't occur free, and (c) prove a type-theoretic version of theorem (471), which asserts $\mathit{Situation}(x) \to ((x \models p) \equiv x[\lambda y\, p])$. We leave these as simple exercises. ⋈

(971.6) (Exercise)

(971.7) From the assumption that $\varphi\downarrow$, the proof now follows that of (511.3). ⋈

(971.9) – **(971.11)** (Exercises)

(971.12) From the assumption that $\varphi\downarrow$, the proof now follows that of (545.5). ⋈

(971.13) (Exercise)

(972.1) Since $[\lambda x_1 \ldots x_n\; w \models F^n x_1 \ldots x_n]$ is a core λ-expression (i.e., none of the variables bound by the λ occur in encoding position), it is axiomatic that $[\lambda x_1 \ldots x_n\, w \models F^n x_1 \ldots x_n]\downarrow$, by (935.5). ⋈

(972.3) (Exercise)

(972.4) Our proof strategy is by conditional proof, outlined as follows:

(A) Assume $A!G$.

(B) Infer $\Box\neg\exists x_1 \ldots \forall x_n G x_1 \ldots x_n$ from (A).

(C) Show: $\Box\neg\exists x_1 \ldots \forall x_n G x_1 \ldots x_n \vdash \forall x_1 \ldots \forall x_n \neg G_w x_1 \ldots x_n$.
(Note that G is in the premise and G_w is in the conclusion.)

(D) Apply Rule RN to (C) to conclude:
$\Box\Box\neg\exists x_1 \ldots \forall x_n G x_1 \ldots x_n \vdash \Box\forall x_1 \ldots \forall x_n \neg G_w x_1 \ldots x_n$.

(E) Apply the 4 schema to (B) to derive $\Box\Box\neg\exists x_1 \ldots \forall x_n G x_1 \ldots x_n$.

(F) Conclude, from (D) and (E) that $\Box\forall x_1 \ldots \forall x_n \neg G_w x_1 \ldots x_n$.

So it remains to show only steps (B) and (C).

(B) Assumption (A) implies $\Box A!_G$, since $A!$ is a rigid property (959.7) [180.2]. Independently, the application of Rule RM to axiom (935.24) yields:

$$\Box A!_G \rightarrow \Box\neg\exists x_1 \ldots \forall x_n G x_1 \ldots x_n$$

Hence, $\Box\neg\exists x_1 \ldots \forall x_n G x_1 \ldots x_n$.

(C) We show this by first establishing:

$$\Box\neg\exists x_1 \ldots \forall x_n G x_1 \ldots x_n \rightarrow \forall x_1 \ldots \forall x_n \neg G_w x_1 \ldots x_n$$

and then applying (937.4) [63.10]. So assume $\Box\neg\exists x_1 \ldots \forall x_n G x_1 \ldots x_n$. It follows by quantified modal logic that $\Box\forall x_1 \ldots \forall x_n \neg G x_1 \ldots x_n$. By n applications of CBF, we therefore know $\forall x_1 \ldots \forall x_n \Box\neg G x_1 \ldots x_n$. Instantiate to arbitrarily chosen objects $x_1, \ldots x_n$, having types $t_1, \ldots, t_n$, respectively, so that we have $\Box\neg G x_1 \ldots x_n$. So by a fundamental theorem of possible worlds (971.9), it follows that $\forall w'(w' \models \neg G x_1 \ldots x_n)$. Instantiate to w, to obtain $w \models \neg G x_1 \ldots x_n$. Then since possible worlds are coherent (971.8), it follows that:

$$\neg w \models G x_1 \ldots x_n$$

But the following is an instance of λ-Conversion:

$$[\lambda x_1 \ldots x_n \, w \models G x_1 \ldots x_n] x_1 \ldots x_n \equiv w \models G x_1 \ldots x_n$$

Our last two results imply:

$$\neg[\lambda x_1 \ldots x_n \, w \models G x_1 \ldots x_n] x_1 \ldots x_n$$

which by definition (972.2) yields $\neg G_w x_1 \ldots x_n$. Since $x_1, \ldots, x_n$ were arbitrarily chosen, it follows that:

$$\forall x_1 \ldots \forall x_n \neg G_w x_1 \ldots x_n \qquad \bowtie$$

(973.2) Assume $\exists z Gz$ and $ExtensionOf(x, G)$. By GEN, we have to show $xF \rightarrow O!_F$. So sssume xF. Then by definition (973.1), $\forall z(Fz \equiv Gz)$. Hence $\exists z Fz$. But then by axiom (935.24), $\neg A!_F$. Then by the type-theoretic counterpart of theorem (222.2) (exercise), $O!_F$. $\bowtie$

(973.3) Assume $\neg\exists z Gz$ and $ExtensionOf(x, G)$. By GEN, we have to show $A!_F \rightarrow xF$. So assume $A!_F$. Then by axiom (935.24), $\neg\exists z Fz$. From this and the assumption that $\neg\exists z Gz$ it follows, by the type-theoretic counterpart of theorem (103.9) (exercise), that $\forall z(Fz \equiv Gz)$. From this, the assumption that $ExtensionOf(x, G)$, and definition (973.1), it follows that xF. $\bowtie$

(973.4) Assume (A) $O!\textsc{g}$, (B) $\neg\exists zGz$, (C) *ExtensionOf*(x,G), (D) $A!\textsc{h}$, and (E) *ExtensionOf*(y,H). By (C), (E), and definition (973.1), we know that $A!x$ and $A!y$. Moreover, since H is abstract (D), we know by axiom (935.24) that $\neg\exists zHz$. From this and (B) it follows that $\forall z(Gz \equiv Hz)$. From this last fact, and the facts that $\forall F(xF \equiv \forall z(Fz \equiv Gz))$ (by (C) and definition 973.1) and $\forall F(yF \equiv \forall z(Fz \equiv Hz))$ (by (E) and definition 973.1), it follows that $\forall F(xF \equiv yF)$. From this fact and the previously established facts $A!x$ and $A!y$, it follows that $x = y$, by (962.2). ⋈

(973.5) (Exercise)

(974.1) – **(974.2)** (Exercises)

(975.3) – **(974.5)** (Exercises)

(978.12) – **(978.16)** (Exercises)

(978.17) Assume *Fictional-R*$(\textsc{f})$. Then if we take the variable x in (978.13) to have type $\langle i\rangle$ and instantiate that variable to $\textsc{f}$, then where $A!$ has type $\langle\langle i\rangle\rangle$, it follows that $A!\textsc{f}$. Note independently that by applying Rule RN to axiom (935.24), we know:

$$\Box(A!\textsc{f} \to \neg\exists xFx)$$

So by the K schema, $\Box A!\textsc{f} \to \Box\neg\exists xFx)$. But we've established $A!\textsc{f}$. So $\Box A!\textsc{f}$, by (959.7) [180.2]. Hence, $\Box\neg\exists xFx$, i.e., $\neg\Diamond\exists xFx$. ⋈

(978.19) (Exercise)

(980.3) We want to show:

$$T = \imath x(A!x \,\&\, \forall F(xF \equiv \exists p(T \models p \,\&\, F = [\lambda y\, p])))$$

Since both terms flanking the identity sign denote situations and thus abstract objects, it suffices to show, by (962.2) and GEN:

$$TG \equiv \imath x(A!x \,\&\, \forall F(xF \equiv \exists p(T \models p \,\&\, F = [\lambda y\, p])))G$$

So, we prove both directions of the biconditional.

$(\to)$ Assume TG, to show:

$$(\vartheta)\quad \imath x(A!x \,\&\, \forall F(xF \equiv \exists p(T \models p \,\&\, F = [\lambda y\, p])))G$$

But note that the description in (ϑ) is *strictly canonical* (963.5) [260.2]. For if we let φ be the formula $\exists p(T \models p \,\&\, F = [\lambda y\, p])$, then it is straightforward to show that:

$$\vdash_\Box \forall F(\varphi \to \Box\varphi)$$

We leave this as an exercise. Then by the modally strict Abstraction Principle proved in (963.5) [261.3], we know that:

$$(\xi)\ \ \iota x(A!x \,\&\, \forall F(xF \equiv \exists p(T \models p \,\&\, F = [\lambda y\, p])))G \equiv \exists p(T \models p \,\&\, G = [\lambda y\, p])$$

So to show (ϑ), it suffices to show $\exists p(T \models p \,\&\, G = [\lambda y\, p])$. Since T is a situation, every property T encodes is propositional (971.2). So G is a propositional property. It follows by (970) that $\exists p(G = [\lambda y\, p])$. Let q be such a proposition, so that we know $G = [\lambda y\, q]$. Since TG by hypothesis, it follows that $T[\lambda y\, q]$. But since T is a situation, this last result implies $T \models q$, by (971.1) and (971.3). So we have established $T \models q \,\&\, G = [\lambda y\, q]$. Hence $\exists p(T \models p \,\&\, G = [\lambda y\, p])$.

($\leftarrow$) Assume $\exists p(T \models p \& G = [\lambda y p])$. Let q be such a proposition, so that we know $T \models q \,\&\, G = [\lambda y\, q]$. Since T is a situation, the first conjunct implies $T[\lambda y\, q]$, by (971.1) and (971.3). But, then by the second conjunct and Rule =E, TG. $\bowtie$

(980.4) (Exercise)

(981.5) To prove contrapositive, i.e., if $\vdash T \models \varphi^*$, then $\vdash_T \varphi$, we first prove a metalemma:

> *Metalemma:* (Contributed by Uri Nodelman): Let T be an imported mathematical theory and let the following be an enumeration of the new axioms introduced by the Importation Principle (981.3):
>
> $$\begin{array}{l} T \models \chi_1^* \\ T \models \chi_2^* \\ \quad\vdots \end{array}$$
>
> Then there are no theorems of the form $T \models \varphi^*$ in which φ^* is distinct from every χ_i^*. That is, for every theorem of the form $T \models \varphi^*$, there must be an i such that $\varphi^* = \chi_i^*$.
>
> *Proof:* It suffices to show that there is a situation $T_{\min}$ such that $T_{\min} \models \chi_i^*$ for all i and such that $\neg T_{\min} \models \varphi^*$ for any φ^* distinct from every χ_i. For if such an object can be constructed consistently by comprehension, then nothing forces additional theorems of the form $T \models \varphi^*$, for φ^* distinct from every χ_i^*. (If the system were to imply additional theorems of that form, then $T_{\min}$ can't be constructed.)
>
> But we can construct such a $T_{\min}$. First, extend object theory with a new statements of the form:
>
> $$\mathit{ClosedTheorem}_T(\psi^*)$$
>
> to signify that ψ is a closed theorem of T. Do a parallel importation so that:

$ClosedTheorem_T(\psi^*)$ is an axiom if and only if $\vdash_T \psi$.

Note that there are no axioms or theorems of our system that allow inferences to new theorems of the form $ClosedTheorem_T(\varphi^*)$ and nothing follows from any statement of this form.

Now, consider the abstract object $T_{\min}$:

$$T_{\min} = \iota x(A!x \,\&\, \forall F(xF \equiv \exists p(ClosedTheorem_T(p) \,\&\, F = [\lambda y\, p])))$$

Note that the axioms imported with the theory T (in the statement of the Lemma) are also true in $T_{\min}$, i.e., we know:

$$\begin{array}{l} T_{\min} \models \chi_1^* \\ T_{\min} \models \chi_2^* \\ \quad\vdots \end{array}$$

But since there are no new theorems of the form $ClosedTheorem_T(\varphi^*)$, there are no theorems of the form $T_{\min} \models \varphi^*$, where φ^* is distinct from every χ_i^*.

Given the existence of $T_{\min}$, it follows that if a set of truths of the form $T \models \chi_i^*$ are axioms in OT and there are no additional axioms of this form, then no additional truth of the form $T \models \varphi^*$, for any φ^* distinct from every χ_i^*, is derivable.

Given this Metalemma, we can now show that $\vdash_T \varphi$ follows from $\vdash T \models \varphi^*$. For if the enumerated list of axioms that result by importing T is $T \models \chi_1^*$, $T \models \chi_2^*$, ..., then by the Metalemma, there are no theorems of the form $T \models \varphi^*$ where φ^* is distinct from every χ_i^*. So there is an i such that $\varphi^* = \chi_i^*$. Suppose k is such an i, so that we know $\varphi^* = \chi_k^*$. Since $T \models \chi_k^*$ is one of the imported axioms, we know that $\vdash_T \chi_k$. But, then, $\vdash_T \varphi$. ⋈

(981.6) Assume:

(ϑ) $\varphi_1, \ldots, \varphi_n \vdash_T \psi$

(ξ_1) $\vdash T \models \varphi_1^*$

$\quad\vdots$

(ξ_n) $\vdash T \models \varphi_n^*$

By (ϑ) and n applications of the Deduction Theorem, we know:

(ω) $\vdash_T \varphi_1 \to (\ldots \to (\varphi_n \to \psi)\ldots)$

Note separately that from (ξ_1) – (ξ_n) we may infer, respectively, by Metarule (981.5), that:

(ζ_1) $\vdash_T \varphi_1$

$\vdots$

(ζ_n) $\vdash_T \varphi_n$

Then by n applications of Modus Ponens in T to (ω) and (ζ_1),...,(ζ_n), it follows that $\vdash_T \psi$. So by the Importation Principle (981.3), $T \models \psi^*$. ⋈

(984.1) We first prove that the condition $T \models F\tau_T$ is a rigid condition on properties F:

> Assume $T \models F\tau_T$. Since T is, by hypothesis, a situation, it follows from (971.3) that $T\Sigma F\tau_T$, and by (971.1) that $T[\lambda y\, F\tau_T]$. Hence by the axiom for the rigidity of encoding (935.30), $\Box T[\lambda y\, F\tau_T]$, i.e., by reversing our definitions, $\Box T \models F\tau_T$. So by conditional proof, $T \models F\tau_T \rightarrow \Box T \models F\tau_T$. Since this is a theorem, it follows by GEN that $\forall F(T \models F\tau_T \rightarrow \Box T \models F\tau_T)$. But we've established this last result by modally strict reasoning. Hence $\vdash_\Box \forall F(T \models F\tau_T \rightarrow \Box T \models F\tau_T)$, by RN. So $T \models F\tau_T$ is a rigid condition on properties.

It follows by (963.5) [261.3] that:

$$\imath x(A!x \,\&\, \forall F(xF \equiv T \models F\tau_T))F \equiv T \models F\tau_T$$

So by the identification of τ_T in (983.1), it follows by Rule =E that:

$$\tau_T F \equiv T \models F\tau_T$$ ⋈

Bibliography

Ackerman, W., 1924, "Begründung des 'tertium non datur' mittels der Hilbertschen Theorie der Widerspruchsfreiheit", *Mathematische Annalen*, 93: 1–36.

Adams, R., 1974, "Theories of Actuality", *Noûs*, 8: 211–231.

Ajdukiewicz, K., 1934a [1978a], "Sprache und Sinne", *Erkenntnis*, 4: 100–138; translated as "Language and Meaning", John Wilkinson (trans.), in J. Giedymin (ed.), *The Scientific World-Perspective and Other Essays (1931–1963)*, Dordrecht: D. Reidel, 1978, pp. 35–66.

Allen, R. E., 1960, "Participation and Predication in Plato's Middle Dialogues", *The Philosophical Review*, 69(2): 147–164.

Anderson, D., and E. Zalta, 2004, "Frege, Boolos, and Logical Objects", *Journal of Philosophical Logic*, 33(1): 1–26.

Armstrong, D., 1989, *A Combinatorial Theory of Possibility*, New York: Cambridge University Press.

Armstrong, D., 1997, *A World of States of Affairs*, New York: Cambridge University Press.

Avigad, J., and R. Zach, 2013, "The Epsilon Calculus", in *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2013/entries/epsilon-calculus/>.

Balaguer, M., 1998, *Platonism and Anti-Platonism in Mathematics*, New York: Oxford University Press.

Barcan, R., 1946, "A Functional Calculus of First Order Based on Strict Implication", *Journal of Symbolic Logic*, 11: 1–16.

Barwise, J., 1985, "The Situation in Logic–III: Situations, Sets, and the Axiom of Foundation", Technical Report No. CSLI-85-26, Stanford: Center for the Study of Language and Information Publications; reprinted in Barwise 1989b, 177–200. (Page reference is to the reprint.)

Barwise, J., 1989a, "Notes on Branch Points in Situation Theory", in Barwise 1989b, 255–276.

Barwise, J., 1989b, *The Situation in Logic*, CSLI Lecture Notes, No. 17, Stanford: Center for the Study of Language and Information Publications.

Barwise, J. and J. Perry, 1980, "The Situation Underground", in J. Barwise and I. Sag (ed.), *Stanford Working Papers in Semantics* (Volume 1), Stanford: Stanford Cognitive Science Group, 1–55.

Barwise, J. and J. Perry, 1981a, "Situations and Attitudes", *Journal of Philosophy*, 78(11): 668–691.

Barwise, J. and J. Perry, 1981b, "Semantic Innocence and Uncompromising Situations", in *Midwest Studies in Philosophy: Volume 6*, P. French, T. Uehling, and H. Wettstein (eds.), Minneapolis: University of Minnesota, 387–404.

Barwise, J. and J. Perry, 1983, *Situations and Attitudes*, Cambridge, MA: MIT Press.

Barwise, J. and J. Perry, 1984, "Shifting Situations and Shaken Attitudes", Technical Report No. CSLI–84–13, Stanford: Center for the Study of Language and Information Publications; reprinted in *Linguistics and Philosophy* 8(1) (1985): 105–161. (Page reference is to the 1984 publication.)

Bealer, G., 1979, "Theories of Properties, Relations, and Propositions", *The Journal of Philosophy*, 76(11): 634–648.

Bealer, G., 1982, *Quality and Concept*, Oxford: Clarendon Press.

Belnap, N., 1993, "On Rigorous Definitions", *Philosophical Studies*, 72(2–3): 115–146.

Berto, F., 2013, "Impossible Worlds", *Stanford Encyclopedia of Philosophy* (Winter 2013 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2013/entries/impossible-worlds/>.

Bishop, E., 1967, *Foundations of Constructive Analysis*, New York: McGraw-Hill.

Boolos, G., 1971, "The Iterative Concept of Set", *The Journal of Philosophy*, 68(8): 215–231; reprinted in Boolos 1998, 13–29.

Boolos, G., 1975, "On Second-Order Logic", *The Journal of Philosophy*, 72: 509–527; reprinted in Boolos 1998, 37–53.

Boolos, G., 1984, "To Be is to Be a Value of a Variable (or to Be Some Values of Some Variables)", *Journal of Philosophy*, 81: 430–449; reprinted in Boolos 1998, 54–72.

Boolos, G., 1985, "Reading the *Begriffsschrift*", *Mind*, 94: 331–344; reprinted in Boolos 1998, 155–170.

Boolos, G., 1986/1987, "Saving Frege From Contradiction", *Proceedings of the Aristotelian Society*, 87: 137–151; reprinted in Boolos 1998, 171–182.

Boolos, G., 1987, "The Consistency of Frege's *Foundations of Arithmetic*", in *On Being and Saying*, J. Thomson (ed.), Cambridge, MA: MIT Press, 3–20; reprinted in Boolos 1998, 183–201.

Boolos, G., 1995, "Frege's Theorem and the Peano Postulates", *The Bulletin of Symbolic Logic*, 1: 317–326; reprinted in Boolos 1998, 291–300. [Page reference is to the reprint.]

Boolos, G., 1996, "On the Proof of Frege's Theorem", in *Benacerraf and His Crit-*

ics, A. Morton and S. Stich (eds.), Cambridge, MA: Wiley-Blackwell, 143–159; reprinted in Boolos 1998, 275–290. [Page reference is to the reprint.]

Boolos, G., 1998, *Logic, Logic, and Logic*, Cambridge, MA: Harvard University Press.

Brandom, R., 2000, *Articulating Reasons: An Introduction to Inferentialism*, Cambridge, MA: Harvard University Press.

Bueno, O., C. Menzel and E. Zalta, 2014, "Worlds and Propositions Set-Free", *Erkenntnis*, 79: 797–820.

Bueno, O., and E. Zalta, 2005, "A Nominalist's Dilemma and its Solution", *Philosophia Mathematica*, 13: 297–307.

Bueno, O., and E. Zalta, 2017, "Object Theory and Modal Meinongianism", *Australasian Journal of Philosophy*, 95(4): 761–778.

Burgess, J., 1984, "Review of C. Wright, *Frege's Conception of Numbers as Objects*", *The Philosophical Review*, 93(4): 638–640.

Buroker, J., 2014, "Port Royal Logic", *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2014/entries/port-royal-logic/>.

Cantor, G., 1895, *Beiträge zur Begründung der transfiniten Mengenlehre*, *Mathematische Annalen*, 46(4): 481–512.

Cantor, G., 1915, *Contributions to the Founding of the Theory of Transfinite Numbers*, translated with an Introduction and Notes by Philip E. B. Jourdain, New York: Dover.

Cantor, G., 1932, *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, Ernst Zermelo (ed.), Berlin: Springer-Verlag.

Carnap, R., 1929, *Abriss der Logistik*, Wien: J. Springer.

Carnap, R., 1947, *Meaning and Necessity*, Chicago: University of Chicago Press.

Carnap, R., 1950, "Empiricism, Semantics, and Ontology", *Revue Internationale de Philosophie*, 4: 20–40; reprinted in R. Carnap, *Meaning and Necessity*, Chicago: University of Chicago Press, second edition, 1956, 205–221. (Page reference is to the reprint.)

Castañeda, H.-N., 1972, "Thinking and the Structure of the World", *Critica*, 6: 43–86; reprinted in *Philosophia*, 4 (1974): 3–40.

Castañeda, H., 1976, "Leibniz's Syllogistico-Propositional Calculus", *Notre Dame Journal of Formal Logic*, 17: 481–500.

Castañeda, H.-N., 1990. "Leibniz's Complete Propositional Logic", *Topoi*, 9(1): 15–28.

Chellas, B., 1980, *Modal Logic: An Introduction*, Cambridge: Cambridge University Press.

Chierchia, G. and R. Turner, 1988, "Semantics and Property Theory", *Linguistics and Philosophy*, 11: 261–302.

Chisholm, R., 1981, *The First Person*, Minneapolis: University of Minnesota

Press.

Church, A., 1932, "A Set of Postulates for the Foundation of Logic", *The Annals of Mathematics* (2nd Series), 33(2): 346–366.

Church, A., 1940, "A Formulation of the Simple Theory of Types", *Journal of Symbolic Logic*, 5(2): 56–68.

Church, A., 1974, "Russellian Simple Type Theory", *Proceedings and Addresses of the American Philosophical Association*, 47: 21–33.

Chwistek, L., 1921, "Antynomje logiki formbalnej", *Przeglad Filozoficzny*, 24: 164–171.

Chwistek, L., 1922, "Über die Antinomien der Prinzipien der Mathematik", *Mathematische Zeitschrift*, 14: 236–243.

Clark, R., 1978, "Not Every Object of Thought Has Being: A Paradox in Naive Predication Theory", *Noûs*, 12: 181–188.

Cocchiarella, N., 1975, "Second-Order Theories of Predication: Old and New Foundations", *Noûs*, 9(1): 33–53.

Cocchiarella, N., 1978, "On the Logic of Nominalized Predicates and its Philosophical Interpretations", *Erkenntnis*, 13: 339–369.

Cocchiarella, N., 1986, "Frege, Russell, and Logicism", in L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized*, Dordrecht: Reidel, 197–252.

Cocchiarella, N., 1988, "Predication Versus Membership in the Distinction Between Logic as Language and Logic as Calculus", *Synthese*, 77: 37–72.

Cocchiarella, N., 1985, "Frege's Double Correlation Thesis and Quine's Set Theories NF and ML", *Journal of Philosophical Logic*, 14(1): 1–39.

Cohen, S. Marc, 1971, "The Logic of the Third Man", *The Philosophical Review*, 80(4): 448–475.

Colyvan, M., and E. Zalta, 1999, "Mathematics: Truth *and* Fiction?", *Philosophia Mathematica*, 7(3): 336–349.

Copeland, B. J., 2002, "The Genesis of Possible World Semantics", *Journal of Philosophical Logic*, 31: 99–137.

Copeland, B. J., 2006, "Meredith, Prior, and the History of Possible Worlds Semantics", *Synthese*, 150: 373–397.

Cook, R., 2016, "Necessity, Necessitism, and Numbers", *Philosophical Forum*, 47(3–4): 385–414.

Cornford, F., 1939, *Plato and Parmenides*, London: Routledge.

Cresswell, M., 1967, "The Interpretation of Some Lewis Systems of Modal Logic", *Australasian Journal of Philosophy*, 45: 198–206.

Cresswell, M., 1972a, "The World is Everything that is the Case", *Journal of Philosophy*, 50: 1–13.

Cresswell, M., 1972b, "Intensional Logics and Logical Truth", *Journal of Philosophical Logic* 1: 2–15.

Curry, H., 1951, *Outlines of a Formalist Philosophy of Mathematics*, Amsterdam:

North Holland.

Curry, H., 1954, "Remarks on the Definition and Nature of Mathematics", *Dialectica*, 8(3): 228-233.

Dean, W., 2020, "Recursive Function", in *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/sum2020/entries/recursive-functions/

Dedekind, R., 1888 [1893, 1939], *Was sind und was sollen die Zahlen*, Braunschweig: Vieweg und Sohn; 1st edition, 1888; 2nd edition, 1893; 7th edition, 1939; translated as 'The Nature and Meaning of Numbers', in W. W. Beman (ed. and trans.), *Essays on the Theory of Numbers*, Open Court Publishing Company: Chicago, 1901; reprinted, New York: Dover, 1963; also translated as 'What Are Numbers and What Should They Be?', H. Pogorzelski, W. Ryan & W. Snyder (eds. and trans.), Orono, ME: Research Institute for Mathematics, 1995.

Demopoulos, W. (ed.), 1995, *Frege's Philosophy of Mathematics*, Cambridge, MA: Harvard University Press.

Divers, J., 2002, *Possible Worlds*, London: Routledge.

Dorr, C., and J. Goodman, 2019, "Diamonds are Forever", *Noûs*, 54(3): 632–665.

Došen, K., 1988, "Second-order Logic Without Variables", in *Categorial Grammar*, W. Buszkowski, W. Marciszewski, and J. van Benthem (eds.), Amsterdam: John Benjamins.

Drake, F., 1974, *Set Theory: An Introduction to Large Cardinals*, Amsterdam: North Holland.

Dudman, V. H., 1973, "Frege on Definition", *Mind*, 83(328): 609–610.

Dummett, M., 1981, *The Interpretation of Frege's Philosophy*, London: Duckworth.

Ebert, P. and M. Rossberg (trans.), 2013, *Basic Laws of Arithmetic*, translation of Frege 1893 and 1903a, with Crispin Wright, Oxford: Oxford University Press, 2013; revised and corrected paperback edition, 2016.

Enderton, H., 1972 [2001], *A Mathematical Introduction to Logic*, San Diego: Academic Press; second edition, 2001.

Erdmann, B., 1892, *Logik*, Halle an der Saale: Niemeyer, 2 volumes.

Evans, G., 1978, "Can There Be Vague Objects", *Analysis*, 38(4): 208.

Ewald, W. B. (ed.), 1996, *From Kant to Hilbert*: *A Source Book in the Foundations of Mathematics* (Volume II), Oxford: Clarendon Press.

Feferman, S., 1995, "Definedness", *Erkenntnis*, 43(3): 295–320.

Findlay, J., 1963 [1933], *Meinong's Theory of Objects and Values*, Oxford: Clarendon Press.

Fine, K., 1977, "Prior on the Construction of Possible Worlds and Instants" (Postscript to A. N. Prior, *Worlds, Times, and Selves*), K. Fine (ed.), London:

University Duckworth, 116–168; reprinted in K. Fine, *Modality and Tense: Philosophical Papers*, Oxford: Oxford University Press, 2005, 133–175.

Fine, K., 1994, "Essence and Modality", *Philosophical Perspectives*, 8: 1–16.

Fitch, G., 1979, "Analyticity and Necessity in Leibniz", *Journal of the History of Philosophy*, 17(1): 29–42; page reference is to the reprint in *Gottfried Wilhelm Leibniz: Critical Assessments* (Volume 1), R. Woolhouse (ed.), London: Routledge, 1994, 290–307.

Fitelson, B., and E. Zalta, 2007, "Steps Toward a Computational Metaphysics", *Journal of Philosophical Logic*, 36(2): 227–247.

Frances, B., 1996, "Plato's Response to the Third Man Argument in the Paradoxical Exercise of the *Parmenides*", *Ancient Philosophy*, 16: 47–64.

Francez, N., and R. Dyckhoff, 2010, "Proof-theoretic Semantics for a Natural Language Fragment", *Linguistics and Philosophy*, 33: 447–477.

Frede, M., 1967, *Prädikation und Existenzaussage*, *Hypomnemata* 18 (Göttingen).

Frede, M., 1992, "Plato's *Sophist* on False Statements", in *The Cambridge Companion to Plato*, R. Kraut (ed.), Cambridge: Cambridge University Press, 397–424.

Frege, G., 1879, *Begriffsschrift: eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*, Halle a. S.: Louis Nebert; translated as *Concept Script, a formal language of pure thought modelled upon that of arithmetic*, by S. Bauer-Mengelberg, in J. van Heijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, Cambridge, MA: Harvard University Press, 1967, 1–82.

Frege, G., 1884, *The Foundations of Arithmetic*, translated by J. L. Austin, Oxford: Blackwell, 1974 (second revised edition).

Frege, G., 1891, "Funktion und Begriff", Jena: Hermann Pohle; translated as "Function and Object", P. Geach (trans.), in *Translations from the Philosophical Writings of Gottlob Frege*, P. Geach and M. Black (eds. and trans.), Oxford: Blackwell, second edition, reprinted 1970, pp. 21–41. (Page reference is to the translation.)

Frege, G., 1892, "Über Begriffe und Gegenstand", *Vierteljahresschrift für wissenschaftliche Philosophie*, 16: 192–205; translated as "On Concept and Object" by P. Geach and M. Black (trans.), in *Mind*, 60(238) (1951); 168–180; reprinted in *Translations from the Philosophical Writings of Gottlob Frege*, P. Geach and M. Black (eds. and trans.), Oxford: Blackwell, second edition, 1970, pp. 42–55. (Page reference is to the 1970 reprint.)

Frege, G., 1893 [2013], *Grundgesetze der Arithmetik*, Volume I, Jena: Verlag Hermann Pohle; reprinted, Hildesheim: Georg Olms Verlagsbuchhandlung, 1962; translated and edited by Philip Ebert and Marcus Rossberg, with Crispin Wright, *Basic Laws of Arithmetic*, Oxford: Oxford University Press, 2013. [All quotes in translation are from Ebert and Rossberg 2013.]

Frege, G., 1903a [2013], *Grundgesetze der Arithmetik*, Volume II, Jena: Verlag Hermann Pohle; reprinted, Hildesheim: Georg Olms Verlagsbuchhandlung, 1962; translated and edited by Philip Ebert and Marcus Rossberg, with Crispin Wright, *Basic Laws of Arithmetic*, Oxford: Oxford University Press, 2013. [All quotes in translation are from Ebert and Rossberg 2013.]

Frege, G., 1903b, "Über die Grundlagen der Geometrie" (First Series), *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 12 (1903): 319–324 (Part I), 368–375 (Part II); translated as "The Foundations of Geometry", by M. E. Szabo, *The Philosophical Review*, 69(1) (1960): 3–17; translation by E.-H. W. Kluge, in *On the Foundations of Geometry and Formal Theories of Arithmetic*, New Haven: Yale University Press, 1971, 22–37, and reprinted in *Collected Papers on Mathematics, Logic and Philosophy*, B. F. McGuinness (ed.), Oxford: Blackwell Publsishers, 1984, 273–284.

Frege, G., 1914, "Über Logik in der Mathematik", unpublished lecture note, in H. Hermes, F. Kambartel, and F. Kaulbach (eds.), *Gottlob Frege: Nachgelassene Schriften* (Band 1), Hamburg: Meiner, 1969; translated as "Logic in mathematics", in H. Hermes, F. Kambartel, and F. Kaulbach (eds.), *Gottlob Frege: Posthumous Writings*, Oxford: Basil Blackwell, 1979, 203–250.

Frege, G., 1980, *Philosophical and Mathematical Correspondence*, G. Gabriel, *et al.* (eds.), H. Kaal (trans.), Chicago: University of Chicago Press.

Furth, M. (trans.), 1967, *The Basic Laws of Arithmetic*, by Gottlob Frege, Berkeley: University of California Press.

Gallin, D., 1975, *Intensional and Higher-Order Modal Logic*: *With Applications to Montague Semantics* (North-Holland Mathematics Studies: Volume 19), Amsterdam: North-Holland.

Geach, P., 1956, "The Third Man Again", *The Philosophical Review*, 65(1): 72–82.

Geach, P., 1976, "Critical Notice", *Mind*, 85 (339): 436–449.

Gödel, K., 1931, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I", *Monatshefte fur Mathematik und Physik*, 38: 173–198.

Gödel, K., 1944, "Russell's Mathematical Logic", in P. Schilpp (ed.), *The Philosophy of Bertrand Russell*, Evanston: Northwestern University, 123–153; reprinted in *Kurt Gödel: Collected Works* (Volume II: Publications 1938–1974), S. Feferman, *et al*. (eds.), Oxford: Oxford University Press, 1990, 119–141.

Goodman, N., 1951, *The Structure of Appearance*, Cambridge, MA: Harvard University Press; 3rd edition, Dordrecht: Reidel, 1977.

Goranko, Valentin and Antje Rumberg, 2022, "Temporal Logic", *The Stanford Encyclopedia of Philosophy* (Summer 2022 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2022/entries/logic-temporal/>.

Gupta, A., 2014, "Definitions", *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2014/entries/definitions/>.

Hazen, A., 1976, "Expressive Completeness in Modal Language", *Journal of Philosophical Logic*, 5: 25–46.

Hazen, A., 1978, "The Eliminability of the Actuality Operator in Propositional Modal Logic", *Notre Dame Journal of Formal Logic*, 19(4): 617–622.

Hazen, A., 1985, "Review of C. Wright, *Frege's Conception of Numbers as Objects*", *Australasian Journal of Philosophy*, 63: 251–254.

Hazen, A., 1990, "Actuality and Quantification", *Notre Dame Journal of Formal Logic*, 31(4): 498–508.

Hazen, A., B. Rin, and K. Wehmeier, 2013, "Actuality in Propositional Modal Logic", *Studia Logica*, 101: 487–503.

Heck, R., 1993, "The Development of Arithmetic in Frege's *Grundgesetze Der Arithmetik*", *Journal of Symbolic Logic*, 58(2): 579–601.

Heck, R., 2011, *Frege's Theorem*, Oxford: Clarendon Press.

Henkin, L., 1950, "Completeness in the theory of types", *Journal of Symbolic Logic*, 15(2): 81–91.

Hilbert, D., 1922, "Neubegründung der Mathematik: Erste Mitteilung", *Abhandlungen aus dem Seminar der Hamburgischen Universität*, 1: 157–177; English translations in Ewald 1996, 1115–1133 and Mancosu 1998, 198–214.

Hilbert, D. and P. Bernays, 1934–1939, *Grundlagen der Mathematik*, 2 volumes (Volume 1: 1934, Volume 2: 1939), Berlin: Springer.

Hintikka, J., 1955, "Form and Content in Quantification Theory", *Acta Philosophica Fennica*, 8: 7–55.

Hintikka, J., 1957, "Modality as Referential Multiplicity", *Ajatus* 20: 49–64.

Hintikka, J., 1959, "Towards a Theory of Definite Descriptions", *Analysis*, 19(4): 79–85.

Hintikka, J., 1975, "Impossible Possible Worlds Vindicated", *Journal of Philosophical Logic*, 4(4): 475–484.

Hodes, H., 1984, "Logicism and the Ontological Commitments of Arithmetic", *The Journal of Philosophy*, 81(3): 123–149.

Hodes, H., 1990, "Where Do the Natural Numbers Come From?", *Synthese*, 84(3): 347–407.

Hodges, W., 2008, "Tarski's Theory of Definitions", in D. Patterson (ed.), *New Essays on Tarski and Philosophy*, Oxford: Oxford University Press, 94–132.

Hughes, G., and M. Cresswell, 1968, *An Introduction to Modal Logic*, London: Methuen; references to the fourth reprint, 1985.

Hughes, G., and M. Cresswell, 1996, *A New Introduction to Modal Logic*, London: Routledge; reprinted 2005. (Page reference is to the 2005 reprint.)

Huntington, E. V., 1933a, "New Sets of Independent Postulates for the Algebra of Logic, with special reference to Whitehead and Russell's *Principia Mathematica*", *Transactions of the American Mathematical Society*, 35(1): 274–304.

Huntington, E. V., 1933b, "Boolean algebra. A Correction", *Transactions of the American Mathematical Society*, 35(2): 557–558.

Ingarden, R., 1931, *Das literarische Kunstwerk: Eine Untersuchung aus dem Grenzgebiet der Ontologie, Logik und Literaturwissenschaft*, Halle: Max Niemeyer; 4th edition, 1972.

Ingarden, R., 1964, *Der Streit um die Existenz Der Welt*, Tübingen: Max Niemeyer.

Ishiguro, H., 1972, *Leibniz's Philosophy of Logic and Language*, London: Duckworth.

Jago, M., 2013, "Impossible Worlds", *Noûs*, 47(3): 713–728.

Jonas, S., 2023, "Mathematical Pluralism and Indispensability", *Erkenntnis*, first online 21 February 2023. doi:10.1007/s10670-023-00659-6

Kant, I., 1787, *The Critique of Pure Reason*, 2nd edition, Norman Kemp-Smith (trans.), New York: St. Martin's Press, 1929.

Kaplan, D., 1970, "S5 With Quantifiable Propositional Variables", *The Journal of Symbolic Logic*, 35(2): 355.

Kaplan, D., 1977 [1989], "Demonstratives", in J. Almog, H. Wettstein, and J. Perry (eds.), *Themes from Kaplan*, Oxford: Oxford University Press, 1989, 481–563.

Kauppi, R., 1960, *Über die Leibnizsche Logik*, Helsinki: Suomalaisen Kirjallisuuden Kirjapaino Oy.

Kauppi, R., 1967, *Einführung in die Theorie der Begriffssysteme* (monograph), in *Acta Universitatis Tamperensis*, Series A, Volume 15, Tampere: Tampereen Yliopisto.

Kim, H., 2022, "A New Class of Fictional Truths", *The Philosophical Quarterly*, 71(1): 90–107.

Kirchner, D., 2017 [2021], *Representation and Partial Automation of the Principia Logico-Metaphysica in Isabelle/HOL*, Masters Thesis, Institute of Mathematics, Freie Universität Berlin, May 2017; updated February 2021 version available online at the Archive of Formal Proofs <https://www.isa-afp.org/browser_info/current/AFP/PLM/document.pdf> Isabelle/HOL input files available online at <https://isa-afp.org/entries/PLM.html>

Kirchner, D., 2022, *Computer-Verified Foundations of Metaphysics and an Ontology of Natural Numbers in Isabelle/HOL*, Ph.D. Dissertation, Fachbereich Mathematik und Informatik der Freien Universität Berlin. doi:10.17169/refubium-35141

Kirchner, D., C. Benzmüller, and E. Zalta, 2020, "Mechanizing *Principia Logico-*

Metaphysica in Functional Type Theory", *Review of Symbolic Logic*, 13(1): 206–18.

Kluge, E.-H. W., 1980, *The Metaphysics of Gottlob Frege: An Essay in Ontological Reconstruction*, Dordrecht: Springer.

Krakauer, B., 2013, "What Are Impossible Worlds?" *Philosophical Studies*, 165(3): 989–1007.

Kripke, S., 1959, "A Completeness Theorem in Modal Logic", *The Journal of Symbolic Logic*, 24(1): 1–14.

Krikpe, S., 1963a, "Semantical Considerations on Modal Logic", *Acta Philosophica Fennica*, 16: 83–94.

Kripke, S., 1963b, "Semantical Analysis of Modal Logic I: Normal Modal Propositional Calculi", *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 9: 67–96.

Kripke, S., 1965, "Semantic Analysis of Modal Logic II: Non-normal Modal Propositional Calculi", in *The Theory of Models* (Proceedings of the 1963 International Symposium at Berkeley), J. Addison, L. Henkin, and A. Tarski (eds.), Amsterdam: North Holland, 206–220.

Kripke, S., 1971, "Identity and Necessity", in M. Munitz (ed.), *Identity and Individuation*, New York: New York University Press, 135–164; reprinted in S. Schwartz (ed.), *Naming, Necessity, and Natural Kinds*, Ithaca: Cornell University Press, 1977, 66–101.

Kripke, S., 1972 [1980], *Naming and Necessity*, in Donald Davidson and Gilbert Harman (eds.), *Semantics of Natural Language*, 253–355; republished with Addenda, Cambridge, MA: Harvard University Press, 1980.

Kripke, S., 1973 [2013], *Reference and Existence: The John Locke Lectures*, Oxford: Oxford University Press, 2013.

Leibniz, G. W. F., 1679, "Elements of a Calculus", in Leibniz *LLP*, 17–24.

Leibniz, G. W. F., 1686, *Discourse on Metaphysics*, in Leibniz *PW*, 18–47; also translated by Jonathan Bennett, 2004 [2007], available online at <http://www.earlymoderntexts.com/assets/pdfs/leibniz1686d.pdf>.

Leibniz, G. W. F., 1690a, "A Study in the Plus-Minus Calculus", in Leibniz *LLP*, 122–130.

Leibniz, G. W. F., 1690b, "A Study in the Calculus of Real Addition", in Leibniz *LLP*, 131–144.

Leibniz, G. W. F., 1709, *Theodicy*, translated by E. Huggard, New Haven: Yale University Press, 1952.

Leibniz, G. W. F., 1714, *The Monadology*, in Leibniz *PW*, 179–194.

Leibniz, G. W. F., [**C**], *Opuscules et fragments inédits de Leibniz*, L. Couturat (ed.), Paris: F. Alcan, 1903.

Leibniz, G. W. F., [**G**], *Die philosophischen Schriften von Gottfried Wilhelm Leibniz*, 7 volumes, C. I. Gerhardt (ed.), Berlin: Weidmannsche Buchhandlung, 1875–90.

Leibniz, G. W. F., [**GI**], *Generales Inquisitiones de Analysi Notionum et Veritatum*, F. Schupp (ed.), Hamburg, 1982.

Leibniz, G. W. F., [**LA**], *The Leibniz-Arnauld Correspondence*, H. T. Mason (trans.), Manchester: Manchester University Press, 1967.

Leibniz, G. W. F., [**LLP**], *Leibniz: Logical Papers*, G. Parkinson (ed. and trans.), Oxford: Clarendon, 1966.

Leibniz, G. W. F., [**PW**], *Philosophical Writings*, G. H. R. Parkinson (ed.), M. Morris and G. Parkinson (trans.), London: Dent & Sons, 1973.

Leinster, T., 2014, "Rethinking Set Theory", *American Mathematical Monthly*, 121(5): 403–415.

Leitgeb, H., 2019, "HYPE: A System of Hyperintensional Logic (with an Application to Semantic Paradoxes)", *Journal of Philosophical Logic*, 48: 305–405.

Leitgeb, H., U. Nodelman, and E. Zalta, m.s., "A Defense of Logicism", preprint available online at <https://mally.stanford.edu/Papers/logicism.pdf>.

Lenzen, W., 1990, *Das System der Leibnizschen Logik*, Berlin: W. de Gruyter.

Leonard, H. S., and N. Goodman, 1940, "The Calculus of Individuals and Its Uses", *Journal of Symbolic Logic*, 5: 45–55.

Lewis, C. I., and C. H. Langford, 1932 [1959], *Symbolic Logic*, New York: Century Company, 1932; second edition, New York: Dover Publications, 1959. (Page reference is to 1959, which contains Appendix III, "Final Note on System S2".)

Lewis, D., 1973, *Counterfactuals*, Oxford: Blackwell. (Page reference is to the 2005 reprint.)

Lewis, D., 1986, *On the Plurality of Worlds*, Oxford: Blackwell.

Lewis, D., 1991, *Parts of Classes*, Oxford: Blackwell.

Linsky, B., 1999, *Russell's Metaphysical Logic*, Stanford: CSLI Publications.

Linsky, B., and E. Zalta, 1994, "In Defense of the Simplest Quantified Modal Logic", *Philosophical Perspectives*, 8: 431–458.

Linsky, B., and E. Zalta, 1995, "Naturalized Platonism vs. Platonized Naturalism", *The Journal of Philosophy*, 92(10): 525–555.

Linsky, B., and E. Zalta, 2006, "What is Neologicism?" *The Bulletin of Symbolic Logic*, 12(1): 60–99.

Lloyd, G., 1978, "Leibniz on Possible Individuals and Possible Worlds", *Australasian Journal of Philosophy*, 56(2): 126–142; page reference is to the reprint in *Gottfried Wilhelm Leibniz: Critical Assessments* (Volume 1), R. Woolhouse (ed.), London: Routledge, 1994, 366–385.

MacBride, F., 2003, "Speaking with Shadows: A Study of Neo-Logicism", *British Journal for the Philosophy of Science*, 54: 103–163.

Maddy, P., 1990, *Realism in Mathematics*, Oxford: Clarendon.

May, R., and K. Wehmeier, 2019, "The Proof of Hume's Principle", in P. Ebert and M. Rossberg, *Essays on Frege's* Basic Laws of Arithmetic, Oxford: Oxford University Press, 182–206.

Mally, E., 1912, *Gegenstandstheoretische Grundlagen der Logik und Logistik*, Leipzig: Barth.

Mancosu, P., 1998, *From Brouwer to Hilbert. The Debate on the Foundations of Mathematics in the 1920s*, Oxford: Oxford University Press.

Mancosu, P., 2017, *Abstraction and Infinity*, Oxford: Oxford University Press.

Mares, E., 1997, "Who's Afraid of Impossible Worlds", *Notre Dame Journal of Formal Logic*, 38(4): 516–526.

Mates, B., 1968, "Leibniz on Possible Worlds", in *Logic, Methodology, and Philosophy of Science III*, B. van Rootselaar and J. Staal (eds.), Amsterdam: North Holland; page references are to the reprint in *Leibniz: A Collection of Critical Essays*, H. Frankfurt (ed.), Notre Dame: Notre Dame University Press, 1976 (Garden City: Anchor Books, 1972).

Mates, B., 1972, *Elementary Logic*, 2nd edition, Oxford: Oxford University Press.

McCune, W., 1997, "Solution of the Robbins Problem", *Journal of Automated Reasoning*, 19(3): 263–276.

McCune, W., R. Veroff, B. Fitelson, K. Harris, A. Feist, and L. Wos, 2002, "Short Single Axioms for Boolean Algebra", *Journal of Automated Reasoning*, 29: 1–16.

McMichael, A., 1983, "A Problem for Actualism about Possible Worlds", *Philosophical Review*, 92: 49–66.

McMichael, A., and E. Zalta, 1980, "An Alternative Theory of Nonexistent Objects", *Journal of Philosophical Logic*, 9: 297–313.

McNamara, P., 1993, "Does the actual world actually exist?" *Philosophical Studies*, 69: 59–81.

Meinong, A., 1904, "*Über Gegenstandstheorie*", in A. Meinong (ed.), *Untersuchungen zur Gegenstandstheorie und Psychologie*, Leipzig: Barth; English translation, "On the Theory of Objects", by I. Levi, B. Terrell, and R. Chisholm, *Realism and the Background of Phenomenology*, R. Chisholm (ed.), Glencoe, IL: The Free Press, 1960, 76–117. (Page reference is to the translation.)

Meinong, A., 1910, *Über Annahmen*, 2nd edition, Leipzig: Barth; English translation, *On Assumptions*, edited and with an Introduction by James Heanue, Berkeley: University of California Press, 1983. (Page reference is to the translation.)

Meinwald, C., 1992, "Goodbye to the Third Man", in *The Cambridge Companion to Plato*, R. Kraut (ed.), Cambridge: Cambridge University Press, 365–396.

Mendelson, E., 1964 [1997], *Introduction to Mathematical Logic*, Boca Raton: Chapman & Hall/CRC Press; fourth edition, 1997.

Menzel, C., 1986, "A Complete, Type-Free 'Second-order' Logic and Its Philo-

sophical Foundations", Technical Report #CSLI-86-40, Stanford: Center for the Study of Language and Information Publications; available at <http://cslipublications.stanford.edu/TechReports/CSLI-86-40.pdf>.
Menzel, C., 1990, "Actualism, Ontological Commitment, and Possible World Semantics", *Synthese*, 85: 355–389.
Menzel, C., 2015, "Possible Worlds", *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2015/entries/possible-worlds/>.
Menzel, Christopher, 2016, "Actualism", *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/actualism/>.
Menzel, C., 2020, "In Defense of the Possibilism-Actualism Distinction", *Philosophical Studies*, 177: 1971–1997.
Menzel, C., and E. Zalta, 2014, "The Fundamental Theorem of World Theory", *Journal of Philosophical Logic*, 43(2): 333–363.
Moltmann, F., 2013, *Abstract Objects and the Semantics of Natural Language*, Oxford: Oxford University Press.
Moltmann, F., 2020, "Abstract Objects and the Core-Periphery Distinction in the Ontological and the Conceptual Domain of Natural Language", in J. L. Falguera and C. Martínez-Vidal (eds.), Cham: Springer, 255–276.
Mondadori, F., 1973, "Reference, Essentialism, and Modality in Leibniz's Metaphysics", *Studia Leibnitiana*, 5(1): 74–101; page reference is to the reprint in *Gottfried Wilhelm Leibniz: Critical Assessments* (Volume 1), R. Woolhouse (ed.), London: Routledge, 1994, 230–255.
Mondadori, F., 1975, "Leibniz and the Doctrine of Inter-World Identity", *Studia Leibnitiana*, 7(1): 21–57; page reference is to the reprint in *Gottfried Wilhelm Leibniz: Critical Assessments* (Volume 1), R. Woolhouse (ed.), London: Routledge, 1994, 256–289.
Montague, R., 1960, "Logical Necessity, Physical Necessity, Ethics, and Quantifiers", *Inquiry*, 4: 259–269; reprinted in *Formal Philosophy: Selected Papers of Richard Montague*, R. Thomason (ed.), New Haven: Yale University Press, 1974.
Montague, R., 1973, "The Proper Treatment of Quantification in Ordinary English", in K. J. J. Hintikka, J. M. E. Moravcsik, and P. Suppes (eds.), *Approaches to Natural Language* (Synthese Library: Volume 49), Dordrecht: D. Reidel, 221–242; reprinted in Montague 1974, 247–270.
Montague, R., 1974, *Formal Philosophy*, R. Thomason (ed.), New Haven: Yale University Press.
Morgan, C., 1973, "Systems of Modal Logic for Impossible Worlds", *Inquiry*, 16: 280–289.
Muskens, R., 1989a, "A Relational Formulation Of The Theory Of Types", *Lin-*

guistics and Philosophy, 12: 325–346.

Muskens, R., 1989b, *Meaning and Partiality*, Ph.D. Thesis, University of Amsterdam.

Myhill, J., 1963, "An Alternative to the the Method of Extension and Intension", in *The Philosophy of Rudolf Carnap*, P. Schilpp (ed.), La Salle: Open Court, 299–310.

Nodelman, U., and E. Zalta, 2014, "Foundations for Mathematical Structuralism", *Mind*, 123(489): 39–78.

Nolan, D., 1997, "Impossible Worlds: A Modest Approach", *Notre Dame Journal of Formal Logic*, 38(4): 535–572.

Nolan, D., 2013, "Impossible Worlds", *Philosophy Compass*, 8(4): 360–372.

Oppenheimer, P. and E. Zalta, 2007, "Reflections on the Logic of the Ontological Argument", *Studia Neoartistotelica*, 4(1): 28–35.

Orey, S., 1959, "Model Theory for the Higher Order Predicate Calculus", *Transactions of the American Mathematical Society*, 92: 72–84.

Padoa, A., 1900, "Logical introduction to any deductive theory", in Jean Van Heijenoort (ed.), *From Frege to Goedel: A Source Book in Mathematical Logic, 1879–1931*, Cambridge, MA: Harvard University Press, 119–123.

Panza, M., 2018, "Was Frege a Logicist for Arithmetic?", in *Eva Picardi on Language, Analysis and History*, Annalisa Coliva, Paolo Leonardi and Sebastiano Moruzzi (eds.), Cham: Palgrave Macmillan, 87–112.

Parsons, C., 1965, "Frege's Theory of Number", *Philosophy in America*, M. Black (ed.), Ithaca: Cornell University Press, 180–203; reprinted with Postscript in Demopoulos 1995, 182–210.

Parsons, T., 1980, *Nonexistent Objects*, New Haven: Yale University Press.

Paśniczek, J., 1994, "Non-standard Possible Worlds, Generalised Quantifiers, and Modal Logic", in *Philosophical Logic in Poland*, J. Wolenski (ed.), Dordrecht: Kluwer, 187–198.

Peano, G., 1889, *Arithmetices Principia: Nova Methodo Exposita*, Torino: Bocca, 1889; translated as "The Principles of Arithmetic, Presented by a New Method", in Jean van Heijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, Cambridge, MA: Harvard University Press, 1967, 83–97.

Pelletier, F. J., and E. Zalta, 2000, "How to Say Goodbye to the Third Man", *Noûs*, 34(2): 165–202.

Plantinga, A., 1974, *The Nature of Necessity*, Oxford: Oxford University Press.

Plantinga, A., 1976, "Actualism and Possible Worlds", *Theoria*, 42: 139–160.

Plato, *The Collected Dialogues of Plato*, E. Hamilton and H. Cairns, Princeton: Princeton University Press, seventh printing, 1973.

Pollock, J., 1967, "The Logic of Logical Necessity", *Logique et Analyse*, 10: 307–323.

Pollock, J., 1984, *The Foundations of Philosophical Semantics*, Princeton: Princeton University Press.

Prawitz, D., 1973, "Towards a Foundation of a General Proof Theory", in P. Suppes, et al. (eds.), *Logic, Methodology, and Philosophy of Science* (Volume IV), Amsterdam: North-Holland, pp. 225–307.

Prawitz, D., 2006, 'Meaning Approached via Proofs', *Synthese*, 148: 507–524.

Priest, G., 1992, "What is a Non-Normal World?" *Logique et Analyse*, 35: 291–302.

Priest, G., 2016, *Towards Non-Being: The Logic and Metaphysics of Intentionality*, second edition, Oxford: Oxford University Press; first edition, 2005.

Priest, G. and R. Sylvan, 1992, "Simplified Semantics for Basic Relevant Logics", *Journal of Philosophical Logic*, 21: 217–232.

Prior, A. N., 1956, "Modality and Quantification in S5", *The Journal of Symbolic Logic* 21(1): 60–62.

Putnam, Hilary, 1967a, "Mathematics Without Foundations," *The Journal of Philosophy*, 64(1): 5–22; reprinted in P. Benacerraf and H. Putnam (eds.), *Philosophy of Mathematics*, Cambridge: Cambridge University Press, 1983, 295–311.

Putnam, Hilary, 1967b, "The Thesis that Mathematics is Logic," in R. Schoenman (ed.), *Bertrand Russell, Philosopher of the Century*, London: Allen and Unwin, 273–303; reprinted in H. Putnam, *Mathematics, Matter, and Method*, Cambridge: Cambridge University Press, 1979, 12–42.

Prior, A. N., 1963, "Possible Worlds", *The Philosophical Quarterly*, 12(46): 36–43.

Quine, W. V. O., 1948, "On What There Is", *Review of Metaphysics*, 2: 21–38; reprinted in *From a Logical Point of View*, New York: Harper, 1953, 1–19.

Quine, W. V. O., 1951, "Two Dogmas of Empiricism", *Philosophical Review*, 60(1): 20–43; reprinted in W. V. O. Quine, *From a Logical Point of View*, second, revised edition, Cambridge, MA: Harvard University Press, 1961, 20–46. (Page reference is to the original.)

Quine, W. V. O., 1953 [1961], "Reference and Modality", in *From a Logical Point of View*, New York: Harper, 2nd edition, 1961, 139–159.

Quine, W. V. O., 1956, "Quantifiers and Propositional Attitudes", *Journal of Philosophy*, 53(5): 177–187; reprinted in *The Ways of Paradox and Other Essays*, revised and enlarged edition, Cambridge, MA: Harvard, 1976, 185–196.

Quine, W. V. O., 1960, "Variables Explained Away", *American Philosophical Society*, 104(3): 343–347; reprinted in *Selected Logical Papers*, New York: Random House, 1966, 227–235.

Quine, W. V. O., 1953 [1966], "Three Grades of Modal Involvement", in *Proceedings of the XIth International Congress of Philosophy* (Brussels, 1953, Volume 14), Amsterdam: North-Holland Publishing Co.; reprinted in *The Ways of Paradox and Other Essays*, New York: Random House, 1966, 156–174.

Quine, W. V. O., 1968, "Propositional Objects", *Critica: Revista Hispanoamericana de Filosofia*, 2(5): 3–29; reprinted in W. V. Quine, *Ontological Relativity and Other Essays*, New York: Columbia University Press, 1969, 139–160.

Quine, W. V. O., 1970, *Philosophy of Logic*, Englewood Cliffs, NJ: Prentice Hall; second edition, Cambridge, MA: Harvard University Press; 1986.

Quine, W. V. O., 1972, "Algebraic Logic and Predicate Functors", in R. Rudner and I. Scheffler (eds.), *Logic and Art: Essays in Honor of Nelson Goodman*, New York: Bobbs-Merrill; reprinted in *The Ways of Paradox and Other Essays* (revised and enlarged edition), Cambridge, MA: Harvard University Press, 1976, 283–307.

Quine, W. V. O., 1981a, "Five Milestones of Empiricism", in *Theories and Things*, Cambridge, MA: Belknap Press, 67–72.

Quine, W. V. 0., 1981b, *Theories and Things*, Cambridge, MA: Harvard University Press.

Quine, W. V. O., 1986, *Philosophy of Logic*, second edition, Cambridge, MA: Harvard University Press.

Ramsey, F. P., 1926, "The Foundations of Mathematics", *Proceedings of the London Mathematical Society* (Series 2), 25: 338–384; reprinted in F. P. Ramsey, *The Foundations of Mathematics and other Logical Essays*, R. B. Braithwaite (ed.), London: Kegan Paul, Trench, Trubner, & Co., 1931, pp. 1–61.

Rantala, V., 1982, "Impossible World Semantics and Logical Omniscience", *Acta Philosophica Fennica*, 35: 106–115.

Rapaport, W., 1976, *Intentionality and the Structure of Existence*, Ann Arbor: University Microfilms, #77-1930.

Rapaport, W., 1978, "Meinongian Theories and a Russellian Paradox", *Noûs*, 12: 153–180.

Reck, E., "Dedekind's Contributions to the Foundations of Mathematics", *The Stanford Encyclopedia of Philosophy* (Summer 2016 Edition), Edward N. Zalta (ed.), URL =
<https://plato.stanford.edu/archives/sum2016/entries/dedekind-foundations/>.

Rescher, N., 1954, "Leibniz's Interpretation of His Logical Calculi", *Journal of Symbolic Logic*, 19(1): 1–13.

Rescher, N. and R. Brandom, 1980, *The Logic of Inconsistency: A Study in Non-Standard Possible World Semantics and Ontology*, Oxford: Basil Blackwell.

Resnik, M., 1980, *Frege and the Philosophy of Mathematics*, Ithaca: Cornell University Press.

Routley, R., 1980, *Exploring Meinong's Jungle and Beyond*, Departmental Monograph #3, Philosophy Department, Research School of Social Sciences, Australian National University, Canberra.

Restall, G., 1997, "Ways Things Can't Be", *Notre Dame Journal of Formal Logic*, 38(4): 583–596.

Routley, R., and V. Routley, 1972, "The Semantics of First Degree Entailment", *Noûs*, 6(4): 335–369.

Russell, B., 1902, "Letter to Frege", in J. van Heijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, Cambridge, MA: Harvard University Press, 1967, 124–125.

Russell, B., 1905, "On Denoting", *Mind*, 14: 479–493.

Russell, B., "On 'Insolubilia' and Their Solution by Symbolic Logic", in B. Russell, *Essays in Analysis*, D. Lackey (ed.), New York: Braziller, 1973, pp. 190–214.

Russell, B., 1908, "Mathematical Logic As Based On the Theory of Types", *American Journal of Mathematics*, 30(3): 222–262.

Schroeder-Heister, P., 2013, "Proof-Theoretic Semantics", *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2013/entries/proof-theoretic-semantics/>

Sellars, W., 1955, "The Third Man", *The Philosophical Review*, 64(3): 405–437.

Sellars, W., 1974, "Meaning as Functional Classification", *Synthese*, 27: 417–437.

Semeijn, M., 2021, *Fiction and Common Ground: A Workspace Account* (Ph.D. Thesis, University of Groningen), Alblasserdam: Ridderprint.

Simons, P., 1987, *Parts. A Study in Ontology*, Oxford: Clarendon Press; reprinted Oxford: Oxford University Press, 2003. All page references are to the reprint.

Simpson, S., 2009, *Subsystems of Second Order Arithmetic*, second edition, Cambridge: Cambridge University Press.

Sosa, E., 1985, "Imagery and Imagination: Sensory Images and Fictional Characters", *Grazer Philosophische Studien*, 25: 485–499.

Spinoza, B., 1677, *Ethics*, in E. Curley (trans.), *The Collected Writings of Spinoza* (Volume 1), Princeton: Princeton University Press, 1985.

Stalnaker, R., 1968, "A Theory of Conditionals", *Studies in Logical Theory* (American Philosophical Quarterly Monograph Series: Volume 2), Oxford: Oxford University Press, 98–112.

Stalnaker, R., 1976, "Possible Worlds", *Noûs*, 10: 65–75.

Stalnaker, R., 1984, *Inquiry*, Cambridge, MA: MIT Press.

Stalnaker, R., 2012, *Mere Possibilities: Metaphysical Foundations for Modal Semantics*, Princeton: Princeton University Press.

Strang, C., 1963, "Plato and the Third Man", *Proceedings of the Aristotelian Society* (Supplementary Volume), 37: 147–164.

Suppes, P., 1957, *Introduction to Logic*, Princeton: Van Nostrand.

Swoyer, C., 1994, "Leibniz's Calculus of Real Addition", *Studia Leibnitiana*, 26(1): 1–30.

Swoyer, C., 1995, "Leibniz on Intension and Extension", *Noûs*, 29(1): 96–114.

Swoyer, C., 1996, "Theories of Properties: From Plenitude to Paucity", *Philosophical Perspectives*, 10: 243–264.

Swoyer, C., 1998, "Complex Predicates and and Logics for Properties and Relations", *Journal of Philosophical Logic*, 27(3): 295–325.

Tarski, A., 1930, "Über einige fundamentale Begriffe der Metamathematik", *Comptes Rendus des s:eances de la Société des Sciences et des Lettres de Varsovie*, 23: 22–29; translated by J. H. Woodger as "On Some Fundamental Concepts of Metamathematics", in Tarski 1983, 30–37. (Page reference is to the translation.)

Tarski, A., 1933, "Pojęcie prawdy w językach nauk dedukcyjnych", *Prace Towarzystwa Naukowego Warszawskiego, Wydzial III Nauk Matematyczno-Fizycznych*, 34, Warsaw; expanded version translated by J. H. Woodger as "The Concept of Truth in the Languages of the Deductive Sciences", in Tarski 1983, 152–278.

Tarski, A., 1944, "The Semantic Conception of Truth", *Philosophy and Phenomenological Research*, 4(3): 341–376.

Tarski, A., 1983, *Logic, Semantics, Metamathematics* (Papers from 1923 to 1938), John Corcoran (ed.), Indianapolis: Hackett Publishing Company.

Taylor, A. E., 1915-1916, "Parmenides, Zeno, and Socrates", *Proceedings of the Aristotelian Society* (New Series), 16: 234–289.

Taylor, A. E., 1934, *The* Parmenides *of Plato*, Oxford: Clarendon Press.

Tegmark, M., 2008, "The Mathematical Universe", *Foundations of Physics*, 38: 101–150. Available at <http://arxiv.org/pdf/0704.0646v2.pdf>.

Tennant, N., 2004, "A General Theory of Abstraction Operators", *The Philosophical Quarterly*, 54(214): 105–133.

Urbaniak, R., and S. Hämäri, 2012, "Busting a Myth about Leśniewski and Definitions", *History and Philosophy of Logic*, 33(2): 159–189.

Vailati, E., 1986, "Leibniz on Necessary and Contingent Predication", *Studia Leibnitiana*, 18(2): 195–210.

van Inwagen, P., 1977, "Creatures of Fiction", *American Philosophical Quarterly*, 14 (4): 299–308.

van Inwagen, P., 1983, "Fiction and Metaphysics", *Philosophy and Literature*, 7(1): 67–78.

van Inwagen, P., 1986, "Two Concepts of Possible Worlds", in *Midwest Studies in Philosophy, XI*, P. French, T. Uehling, and H. Wettstein (eds.), Minneapolis: University of Minnesota Press, 185–213.

van Inwagen, P., 2003, "Existence, Ontological Commitment and Fictional Entities", in *The Oxford Handbook of Metaphysics*, M. Loux and D. Zimmerman (eds.), Oxford: Oxford University Press, 131–158.

Varzi, A., 2015, "Mereology", *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2015/entries/mereology/>.

Vlastos, G., 1954, "The Third Man Argument in the Parmenides", *The Philosophical Review*, 63(3): 319–349.

Vlastos, G., 1955, "Addenda to the Third Man Argument: A Reply to Professor Sellars", *The Philosophical Review*, 64(3): 438–448.

Vlastos, G., 1956, "Postscript to the Third Man: A Reply to Mr. Geach", *The Philosophical Review*, 65(1): 83–94.

Vlastos, G., 1969, "Plato's 'Third Man' Argument (PARM. 132A1–B2): Text and Logic", *The Philosophical Quarterly*, 19(77): 289–301.

von Neumann, John, 1923, "Zur Einführung der transfiniten Zahlen", *Acta litterarum ac scientiarum Ragiae Universitatis Hungaricae Francisco-Josephinae* (Sectio scientiarum mathematicarum), 1: 199–208; "On the Introduction of Transfinite Numbers", in Jean van Heijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, 3rd edition, Cambridge, MA: Harvard University Press, pp. 346–354.

von Neumann, J., 1931, "The Formalist Foundations of Mathematics", in A. Taub (ed.), *John von Neumann: Collected Works* (Volume 2), Pergamon Press, New York, 1961.

Wang, H., 1974, *From Mathematics to Philosophy*, New York: Humanities Press

Wang, H., 1996, *A Logical Journey: From Gödel to Philosophy*, Cambridge: MIT Press.

Weber, H., 1893, "Leopold Kronecker", in *Mathematische Annalen*, 43(1) (1893): 1–25; reprinted from *Jahresberichte der Deutschen Mathematiker-Vereinigung*, Band 2, 1892, pp. 5–31.

West, D., m.s., "Foundational Choices for Tensed Object Theory", June 2022, unpublished.

Wetzel, L., 2009, *Types and Tokens: On Abstract Objects*, Cambridge, MA: MIT Press.

Whitehead, A. N., and B. Russell, 1910–1913 [1925–1927], *Principia Mathematica*, 3 volumes, Cambridge: Cambridge University Press; 2nd edition, Cambridge: Cambridge University Press, 1925–1927. (Proposition and page numbers refer to the 2nd edition.)

Williamson, T., 2013, *Modal Logic as Metaphysics*, Oxford: Oxford University Press.

Wilson, M., 1979, "Possible Gods", *Review of Metaphysics*, 32(4): 717–733.

Wittgenstein, L., 1921, *Logisch-philosophische Abhandlung*, translated as *Tractatus Logico-Philosophicus*, by D. F. Pears and B. F. McGuinness, London:

Routledge & Kegan Paul, 1961; 4th impression, 1969. (Page reference is to the 1969 edition.)

Wittgenstein, L., 1953, *Philosophical Investigations*, G. E. M. Anscombe and R. Rhees (eds.), G. E. M. Anscombe (trans.), Oxford: Blackwell.

Wright, C., 1983, *Frege's Conception of Numbers as Objects*, Aberdeen: Aberdeen University Press.

Yagisawa, T., 1988, "Beyond Possible Worlds", *Philosophical Studies*, 53: 175–204.

Zalta, E., 1983, *Abstract Objects: An Introduction to Axiomatic Metaphysics*, Dordrecht: D. Reidel.

———, 1987, "On the Structural Similarities Between Worlds and Times", *Philosophical Studies*, 51(2): 213–239.

———, 1988a, *Intensional Logic and the Metaphysics of Intentionality*, Cambridge, MA: MIT Press.

———, 1988b, "Logical and Analytic Truths That Are Not Necessary", *The Journal of Philosophy*, 85(2): 57–74.

———, 1991, "A Theory of Situations", in *Situation Theory and Its Applications*, J. Barwise, J. Gawron, G. Plotkin, and S. Tutiya (eds.), Stanford: Center for the Study of Language and Information Publications, 81–111.

———, 1993, "Twenty-Five Basic Theorems in Situation and World Theory", *Journal of Philosophical Logic*, 22: 385–428.

———, 1997a, "A Classically-Based Theory of Impossible Worlds", *Notre Dame Journal of Formal Logic*, 38(4): 640–660.

———, 1997b, "The Modal Object Calculus and Its Interpretation", *Advances in Intensional Logic*, Maarten de Rijke (ed.), Dordrecht: Kluwer, 249–279.

———, 1998, "Mally's Determinates and Husserl's Noemata", in *Ernst Mally—Versuch einer Neubewertung*, A. Hieke (ed.), St. Augustin: Academia-Verlag, 9–28.

———, 1999, "Natural Cardinals and Natural Numbers as Abstract Objects: A Partial Reconstruction of Frege's *Grundgesetze* in Object Theory", *Journal of Philosophical Logic*, 28(6): 619–660.

———, 2000a, "A (Leibnizian) Theory of Concepts", *Logical Analysis and History of Philosophy/Philosophiegeschichte und logische Analyse*, 3: 137–183.

———, 2000b, "Neologicism? An Ontological Reduction of Mathematics to Metaphysics", *Erkenntnis*, 53(1–2): 219–265.

———, 2000c, "The Road Between Pretense Theory and Object Theory", in *Empty Names, Fiction, and the Puzzles of Non-Existence*, A. Everett and T. Hofweber (eds.), Stanford: CSLI Publications, 2000, 117–147.

———, 2004, "In Defense of the Law of Noncontradiction", in *The Law of Non-contradiction: New Philosophical Essays*, G. Priest, JC Beall, and B. Armour-Garb (eds.), Oxford: Oxford University Press, 418–436.

———, 2006a, “Deriving and Validating Kripkean Claims Using the Theory of Abstract Objects”, *Noûs*, 40(4): 591–622.

———, 2006b, “Essence and Modality”, *Mind*, 115(459): 659–693.

———, 2014, “The Tarski T-Schema is a Tautology (Literally)”, *Analysis*, 74(1): 5–11.

———, 2017, “Frege’s Theorem and Foundations for Arithmetic”, *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/frege-theorem/>.

———, 2020, “Typed Object Theory”, *Abstract Objects: For and Against*, J. L. Falguera and C. Martínez-Vidal (eds.), Cham: Springer, 59–88.