A Defense of Logicism*

Hannes Leitgeb
Munich Center for Mathematical Philosophy
Ludwig-Maximilians Universität München
hannes.leitgeb@lmu.de

Uri Nodelman
Philosophy Department
Stanford University
nodelman@stanford.edu

Edward N. Zalta
Philosophy Department
Stanford University
zalta@stanford.edu

Abstract

We argue that logicism, the thesis that mathematics is reducible to logic and analytic truths, is true. We do so by (a) developing a formal framework with comprehension and abstraction principles, (b) giving reasons for thinking that this framework is part of *logic*, (c) showing how the denotations for terms and predicates of an arbitrary mathematical theory can be viewed as logical objects that exist in the framework, and (d) showing how each theorem of a mathematical theory can be given a true reading in the logical framework.

*This paper originated as a presentation that the third author prepared for the 31st Wittgenstein Symposium, in Kirchberg, Austria, August 2008. Discussions between the co-authors after this presentation led to a collaboration on, and further development of, the thesis and the technical material grounding the thesis. The authors would especially like to thank Daniel Kirchner for suggesting important refinements of the technical development. We'd also like to thank Allen Hazen, Bernard Linsky, and Otávio Bueno for comments on our argument.

In this paper, we defend logicism, i.e., the claim that mathematics is reducible to logical and analytic truths alone, in the sense that the axioms and theorems of mathematics are derivable from logical truths and analytic truths. We shall assume, in what follows, that the deductive system of second-order logic is a part of logic, both in the usual contemporary sense of *logic* but also in the sense of logic developed later in the paper. This assumption doesn't require a second-order, logical consequence relation, and so our assumption about second-order logic doesn't require any set theory.

In the following defense of logicism, we reinterpret the formalism developed in Zalta 2000 and in Linsky & Zalta 2006, and develop reasons for thinking that this formalism is part of logic. In those previous papers, it was assumed that the formalism was metaphysical in character, and so those papers assumed that logicism, traditionally conceived, was a 'non-starter' and that some form of neologicism was therefore the best one can achieve by way of a desirable fallback position.

More specifically, in those earlier papers, the authors presupposed a standard notion of logical truth and argued that logicism couldn't be true because (i) mathematical theories are often committed to a large, often infinite, ontology, (ii) logic, understood to include second-order logic, is committed only to a non-empty domain of individuals and a 2-element domain of properties, and (iii) the standard for reducing mathematics to logic is relative interpretability. Given these facts, there is no way to reduce the axioms of mathematical theories that have strong existence assumptions to theorems of logic.

In what follows, however, we argue that logicism is true, and indeed, that it can be given a serious defense. Our defense is based on a more nuanced notion of logical truth. Since the notion of logical truth defined in what follows yields a new body of such truths, this leads us to revise both (ii) and (iii) above. If logic is constituted by our new body of logical truths, then contrary to (ii), logic is committed to more than to a non-empty domain of individuals and a 2-element domain of properties; indeed, it may be committed to much more. Contrary to (iii), both the conceptual and epistemological goals of the logicists can be achieved by adopting a notion of reduction other than relative interpretability. We suggest that relative interpretability is the *wrong* notion of reduction and suggest an alternative. We argue that this alternative notion of reduction gives up nothing important when it comes to establishing the most important goals set by the logicists for the foundations of mathematics.

In our defense of logicism, we shall attempt to show that logic includes special domains of individuals, properties, and relations, all of which can be asserted to exist by logical axioms. (Henceforth, we use 'objects' to refer generally to individuals, properties, and relations.) Thus, we agree with the early logicists that logic does have its own special logical objects. But we plan to *justify* this assumption in Section 5, when we defend logicism.

Moreover, when we add certain analytic truths to our background system, we'll be able to assert the existence of new logical objects. Given these new objects, our revised notion of reduction should be familiar: (a) every well-defined individual term of a mathematical theory T is assigned a logical individual as its denotation, (b) every well-defined property or relation term of T is assigned a logical property or logical relation as its denotation, and (c) every theorem of T is assigned a reading stated in terms of these denotations on which it turns out true. Thus, we provide precise theoretical descriptions of the entities denoted by the singular terms and predicates of mathematical theories and this provides the means of stating precise truth conditions of the theorems and non-theorems of mathematical theories. So once we establish that our background system is part of logic and that we've only extended it with *bona fide* analytic truths, we will be defending logicism with respect to a genuine notion of reduction. And with a genuine reduction of mathematics to logic, we achieve the philosophical goals that were foremost in the minds of the early logicists.

1 The Goals of Logicism

Why did logicians and philosophers in the very early 20th century, such as Frege (1893/1903) and Whitehead & Russell (1910–1913), set out to establish the logicist thesis that mathematics is reducible to the laws of logic and analytic truths? If logicism were true, what would be the philosophical benefits?

We take it that there are both conceptual and epistemological benefits. The conceptual benefit is clear: if mathematics is reducible to logic, the conceptual machineries of two *a priori* sciences are reduced to one. The concepts of mathematics become nothing other than concepts of logic. This simplifies the philosophy of mathematics, since (a) logicism would

provide an account of all of mathematics, and not just the mathematics that is applied in, or is indispensable for, the natural sciences, and (b) logicism would provide an account of mathematics whether or not the mathematicians conclude that there is only one distinguished, true mathematical theory.

As to the epistemological benefits of logicism, Benacerraf provides one classic formulation:

> But in reply to Kant, logicists claimed that these [mathematical] propositions are a priori because they are analytic—because they are true (false) merely 'in virtue of' the meanings of the terms in which they are cast. Thus to know their meanings is to know all that is required for a knowledge of their truth. No empirical investigation is needed. The philosophical point of establishing the view was nakedly epistemological: logicism, if it could be established, would show that our knowledge of mathematics could be accounted for by whatever would account for our knowledge of language. And, of course, it was assumed that knowledge of language could *itself* be accounted for in ways consistent with empiricist principles, that language was itself entirely learned. Thus, following Hume, all our knowledge could once more be seen as concerning either 'relations of ideas' (analytic and a priori) or 'matters of fact'.
>
> (Benacerraf 1981, 42–43)

So if logical truths are analytic, and mathematics is reducible to logical and other analytic truths, then we would have an explanation of mathematical knowledge.

This doesn't, strictly speaking, rule out the idea that some special faculty of intuition plays a role in our knowledge of mathematics, but only that if there is such a faculty, it is epistemologically innocent, in the sense that it doesn't require that there be a causal mechanism by which abstract mathematical objects give rise to intuitions. We can avoid Gödel's (1964, 268) talk of the analogy with sense perception, but keep the notion of intuition in an enlightened sense. The thesis that intuition provides some means of non-conceptual access to mathematical objects is perfectly consistent with the view that we will be developing here as long as this access is not meant to be causal.

2 The Background Logical Framework

In this section we sketch a *logical framework* though we shall not argue that it is logical until Section 5. This logical framework is needed to understand and ground the analysis of mathematics presented in Section 3. It is based on the logic of encoding described in Zalta 1983, 1988, and in such articles as Zalta 2000, 2006. Readers familiar with these works may wish to skip ahead to the next section. Such readers should be aware that the theory uses the predicates $O!$ and $A!$ to distinguish between *ordinary* and *abstract* objects, where for the purpose of this paper we will take the predicate 'abstract' to be a primitive term of our theory. In a nutshell, abstract entities are individuated by a group of *encoded* properties that they objectify. We think of that understanding of abstractness as overlapping with common usage without coinciding with it exactly, and we are happy to regard 'abstract' as a technical term the meaning of which is given more precisely by AXIOMS 3–7 (in Section 3.3) of our theory.

Our logical framework, in its full generality, is developed within a relational type theory. However, after we present the framework, we'll focus only on a certain fragment. To keep the presentation simple, our analysis will focus on those first- or second-order mathematical theories statable in terms of primitive individual constants and primitive 1- and 2-place predicates. Examples of such mathematical theories include Zermelo-Fraenkel set theory (ZF), Peano Arithmetic (PA), real number theory ($\Re$), etc.[1] We'll therefore motivate our general logical framework with a discussion of these theories. We note that *all* of the philosophically relevant ideas concerning our analysis of mathematics can be understood by examining these basic theories, and that it should be clear how to extend the framework to analyze mathematical theories requiring more expressive power. That is, the logical framework defined later in this section can be further developed in a variety of ways, e.g., to analyze mathematical theories stated in terms of n-place predicates ($n \geq 0$) and not just 1- and 2-place predicates, to allow for function terms and definite descriptions, etc.

[1] We'll assume, for the present purposes, that any functional terms used in the statement of the axioms of these theories have been replaced by predicates and the relevant existence and uniqueness claims.

2.1 Illustrative Examples

We begin by illustrating what our background logical framework must accomplish. We shall take the basic data of mathematics to be *contextualized* mathematical claims, in the same way that sentences with theoretical terms in empirical science only have meaning in the context of scientific theories. For example, a set theorist, given the context of some set theory, might make the following statement:

> No set is an element of the empty set.

The set theory that forms the context of this statement fixes, at least partially, the meaning of the terms that are used in these sentences. So we shall take the above sentence uttered by the mathematician to be short for the following, where T is some set theory:

> In set theory T, no set is an element of the empty set.

The sentence displayed above would typically be represented formally as follows, where $\vdash_T$ indicates theoremhood with respect to theory T and 'S' denotes the property of being a set (relative to T) and '$\emptyset$' is a constant of T that denotes the empty set:

$$\vdash_T \neg\exists y(Sy \;\&\; y \in \emptyset)$$

Now to be even more specific, suppose the theory T in question is Zermelo-Fraenkel set theory, formulated with the primitive constant $\emptyset$ and the primitive 2-place relation $\in$. On this formulation, the fact that the Null Set Axiom is a theorem of ZF is expressed as $\vdash_{\mathrm{ZF}} \neg\exists y(y \in \emptyset)$ instead of as $\vdash_{\mathrm{ZF}} \exists x\neg\exists y(y \in x)$.

Now to analyze ZF, PA, $\Re$, etc., we shall represent their languages within our (higher-order) framework and so include λ-expressions such $[\lambda x\, \varphi]$, $[\lambda F\, \varphi]$, and $[\lambda R\, \varphi]$, all of which are governed by β-Conversion (and α-Conversion) – see Section 3.2.2 below. In $[\lambda x\, \varphi]$, the λ binds the individual variable 'x' to produce an expression that denotes a property of individuals; in $[\lambda F\, \varphi]$, the λ binds the first-level property variable 'F' to produce an expression that denotes a property of first-level properties; and in $[\lambda R\, \varphi]$, the λ binds the first-level 2-place relation variable 'R' to produce an expression that denotes a property of first-level relations. So where '$\emptyset$' denotes the empty set of ZF, 'S' denotes the ZF property of being a set, and $\in$ denotes the membership relation of ZF, we may infer the following sentences from the above theorem understood now as a theorem of ZF (in which the font sizes of the symbols '$\emptyset$', 'S' and '$\in$' are reduced when they are in argument position):[2]

$$\vdash_{\mathrm{ZF}} [\lambda x\, \neg\exists y(Sy \,\&\, y \in x)]\emptyset$$

$$\vdash_{\mathrm{ZF}} [\lambda F\, \neg\exists y(Fy \,\&\, y \in \emptyset)]S$$

$$\vdash_{\mathrm{ZF}} [\lambda R\, \neg\exists y(Sy \,\&\, y\, R\, \emptyset)]\in$$

That is, from the fact that it is a theorem of ZF that no set is an element of the empty set, we know that: (a) it is a theorem of ZF that the empty set exemplifies the (first-level) property of being an (individual) x such that no set is a member of x; (b) it is a theorem of ZF that the property of being a set exemplifies the second-level property of being a property F such that nothing exemplifying F is an element of the empty set; and (c) it is a theorem of ZF that the membership relation exemplifies the second-level property of being a relation R such that no set bears R to $\emptyset$.

Thus, from the single theorem $\exists y(Sy \,\&\, y \in \emptyset)$, we have inferred additional theorems about the properties exemplified by the objects S, $\in$, and $\emptyset$. We shall import *all* of these theorems into our logical framework as *analytic* truths about what is true in ZF. In particular, sentences very much like the following will be analytic truths of our background theory:

(A) $\mathrm{ZF} \models \neg\exists y(Sy \,\&\, y \in \emptyset)$

(B) $\mathrm{ZF} \models [\lambda x\, \neg\exists y(Sy \,\&\, y \in x)]\emptyset$

(C) $\mathrm{ZF} \models [\lambda F\, \neg\exists y(Fy \,\&\, y \in \emptyset)]S$

(D) $\mathrm{ZF} \models [\lambda R\, \neg\exists y(Sy \,\&\, y\, R\, \emptyset)]\in$

These statements have the form $z \models p$, in which 'z' is an individual variable and 'p' is a variable for a proposition. Statements of this form will be explicitly defined in terms of one of the new primitive logical notions embedded within our logical framework. We shall introduce that definition below, but to complete our examples, notice that if we continue to use 'p' as a variable for propositions, 'F' as a variable ranging over first-level properties, and use '$\mathcal{F}$' as a variable ranging over second-level properties of first-level properties, and '$\mathcal{R}$' as a variable ranging over second-level properties of first-level relations, then:

(A) can be obtained by substituting the proposition $\neg\exists y(Sy \,\&\, y \in \emptyset)$ for p in $\mathrm{ZF} \models p$,

(B) can be obtained by substituting $[\lambda x\, \neg\exists y(Sy \,\&\, y \in x)]$ for F in $\mathrm{ZF} \models F\emptyset$,

(C) can be obtained by substituting $[\lambda F\, \neg\exists y(Fy \,\&\, y \in \emptyset)]$ for $\mathcal{F}$ in $\mathrm{ZF} \models \mathcal{F}S$, and

(D) can be obtained by substituting $[\lambda R\, \neg\exists y(Sy \,\&\, y\, R\, \emptyset)]$ for $\mathcal{R}$ in $\mathrm{ZF} \models \mathcal{R}\in$

Consider that we can now, as a matter of logic, single out all and only those first-level properties F that satisfy the open formula $\mathrm{ZF} \models F\emptyset$; single out all and only those second-level properties $\mathcal{F}$ that satisfy the open formula $\mathrm{ZF} \models \mathcal{F}S$; and single out all and only those second-level properties $\mathcal{R}$ that satisfy the open formula $\mathrm{ZF} \models \mathcal{R}\in$.

Now suppose that we can logically objectify each of the groups of properties singled out by these open formulas. To see how, consider the open formula $\mathrm{ZF} \models F\emptyset$. Suppose that the abstract individuals of object theory are in fact logical individuals that intuitively *code up* all and only the first-level properties of individuals satisfying some arbitrary formula φ. In particular, suppose that there is a unique abstract individual that can *code up* all of the first-level properties F such that $\mathrm{ZF} \models F\emptyset$ and that this individual is a logical object (later, we'll argue for this). Using '$A!$' to denote the first-level property of being abstract, and 'xF' to assert that the individual x *encodes* the property F, and definite descriptions of the form $\imath x\varphi$, we could then formulate the following theoretical identification:

$$\emptyset_{\mathrm{ZF}} = \imath x(A!x \,\&\, \forall F(xF \equiv \mathrm{ZF} \models F\emptyset_{\mathrm{ZF}}))$$

> The empty set of the mathematical theory ZF is the abstract individual x that encodes all and only those (first-level) properties F such that in the theory ZF, the ZF empty set exemplifies F.

Here we are deploying the primitive notion of encoding, x *encodes* F, represented by the formula xF, in which the argument term x is written to

[2]In the following examples, we preserve the infix notation for the relation $\in$ by using a formula of the form yRx. However, when we define our logical framework, we will define relational predications in the usual way as having the form Ryx, and the infix variant yRx will be an abbreviation of the former; it is useful for those cases of relation terms such as $\in$ which traditionally appear using infix notation.

the left of the 1-place relation term F. Formulas of the form xF are to be distinguished from the traditional form of n-place exemplification predication $F^n x_1 \ldots x_n$. The logic of encoding has been described in Zalta 1983, 1988, and elsewhere. Encoding is a *primitive* mode of predication that holds between a (hyperintensional) abstract object and the intensional properties by which we conceive of it.[3] Though encoding can't be defined, it can be axiomatized, and we shall review the axioms governing it below. Thus, in the above example, $\emptyset_{\rm ZF}$ is the hyperintensional object that encodes all and only the properties by which it is theoretically conceived, namely, all and only those Fs such that in the theory ZF, the empty set exemplifies F. In Section 5, we plan to show that this abstract, hyperintensional object is in fact a logical object.

Now to extend these ideas to higher types, consider another open formula mentioned above, namely, ZF $\models \mathcal{F}S$. Suppose that there are special first-level *abstract* properties that can *code up* all and only the second-level properties of properties satisfying some open formula φ. In particular, suppose that there is a unique first-level abstract property that can *code up* all of the second-level properties $\mathcal{F}$ that the property of being a set (S) exemplifies in ZF. Using '$A!$' now to denote the second-level property of being abstract, and $F\mathcal{F}$ to assert that the first-level property F *encodes* the second-level property $\mathcal{F}$, we could then formulate the following theoretical identification:[4]

$$S_{\rm ZF} = \iota F(A!F \,\&\, \forall \mathcal{F}(F\mathcal{F} \equiv {\rm ZF} \models \mathcal{F}S_{\rm ZF}))$$

[3]For the purposes of this paper, we are going to assume that properties may in fact be distinct even though they are necessarily equivalent. Traditionally, philosophers have called such distinct-but-necessarily-equivalent properties *hyperintensional*, since they are more fine-grained than functions from worlds to sets of individuals. But we're going to regard all properties, including those that are distinct but necessarily equivalent, as *intensional*, since we are supposing that properties, right from the start, are primitive entities more fine-grained than functions from worlds to sets of individuals. Since abstract objects are correlated (intuitively) with sets of properties, they become hyperintensional: for any group of intensional properties that are expressible in the language with an open formula φ (with variable F free), there is an abstract object that encodes those properties. It is also worth observing here that our model of object theory in the Appendix is a purely extensional model. That, of course, is not the *intended* model of the theory. But an extensional model is sufficient for establishing that our formal system is consistent.

[4]Notice that in the encoding formula $F\mathcal{F}$, we've made the italic 'F' slightly smaller in size, so as to make it clear that F is the argument and $\mathcal{F}$ is the second-level property it encodes.

> The ZF-property of being a set is the (first-level) abstract property F that encodes all and only those second-level properties $\mathcal{F}$ of first-level properties such that in ZF, the ZF-property of being a set exemplifies $\mathcal{F}$.

Clearly, one of the second-level properties encoded by $S_{\rm ZF}$ is the property $[\lambda F\, \neg\exists y(Fy \,\&\, y \in \emptyset)]$.

Finally, consider the last of the open formulas mentioned above, namely, ZF $\models \mathcal{R}\in$. Suppose that there are special first-level abstract relations that can *code up* all and only the second-level properties of relations satisfying some open formula φ. In particular, suppose that there is a unique first-level abstract relation that can *code up* all of the second-level properties $\mathcal{R}$ that the membership relation ($\in$) exemplifies in ZF. Using '$A!$' to denote the second-level property of being abstract, and '$R\mathcal{R}$' to assert that the first-level relation R *encodes* the second-level property $\mathcal{R}$, we could then formulate the following theoretical identification:[5]

$$\in_{\rm ZF} \;=\; \iota R(A!R \,\&\, \forall \mathcal{R}(R\mathcal{R} \equiv {\rm ZF} \models \mathcal{R}\in_{\rm ZF}))$$

> The membership relation of ZF is the first-level abstract relation R that encodes all and only those second-level properties $\mathcal{R}$ of first-level relations such that in the theory ZF, the ZF-membership relation exemplifies $\mathcal{R}$.

Again, in Section 5, we plan to show that these abstract properties and abstract relations are *logical* properties and *logical* relations.

As we shall see, theoretical identifications like the ones described above are an essential component of our reduction of mathematics to logic. It is important here not to regard these theoretical identifications as *definitions* of the expressions on the left-side of the identity sign, for they appear on the right-side as well. Instead, they are to be regarded as theoretical principles of object theory. We are supposing that from a well-defined body of data, i.e., a body of analytic truths of form "In theory T, p", one can 'abstract out' objects that encode all and only the theoretical properties of the individuals and relations denoted by the constants and 1- and 2-place predicates of T. The other essential component of our reduction will be to show how each theorem of T is given a reading on which it is true.

[5]Again, in the encoding formula $R\mathcal{R}$, we've made the italic R slightly smaller in size, so as to make it clear that R is the argument and $\mathcal{R}$ is the second-level property it encodes.

2.2 Definition of the Logical Framework

Our logical framework has to be defined so that the foregoing formal representations are well-formed. We therefore start with a relational type theory, so that we can quantify over objects of higher type. To be specific, let us define a *type* as follows:

> i is a type.
>
> Whenever $t_1, \ldots, t_n$ are any types ($n \geq 0$), $\langle t_1, \ldots, t_n \rangle$ is a type.

We use i as the type for *individuals*, and $\langle t_1, \ldots, t_n \rangle$ as the type for *relations* among objects having types $t_1, \ldots, t_n$, respectively. Henceforth, where t is any type and $n = 1$, we call entities of type $\langle t \rangle$ *properties*. When $n = 0$, we say that $\langle \, \rangle$ is the type for *propositions*. So properties are 1-place relations and propositions are 0-place relations. We continue to use 'object' to refer to entities of any type.

It should be clear that the examples discussed in the previous subsection employed distinguished expressions of the following types:

- expressions of type i denoting individuals
- expressions of type $\langle \, \rangle$ denoting propositions
- expressions of type $\langle i \rangle$ and $\langle i, i \rangle$ denoting first-level unary relations (= first-level properties) and first-level binary relations, respectively.
- expressions of type $\langle \langle i \rangle \rangle$ and $\langle \langle i, i \rangle \rangle$, denoting second-level properties of properties and second-level properties of relations, respectively.

However, we shall henceforth suppose that for every type t, there is a denumerable list of constants and variables for that type. Among the constants of type $\langle t \rangle$, for any type t, we include the distinguished predicate A, which denotes a primitive property of objects of type t, namely, *being abstract*.

Now in order to state the axioms of our logical framework, we define the language $\mathcal{L}$ by (simultaneously) defining the formulas and terms that constitute the well-formed expressions of $\mathcal{L}$.[6] In these definitions, we'll see that whereas the *terms* are either *simple* or *complex*, there are three basic kinds of formulas: *exemplification* formulas, *encoding* formulas, and *complex* formulas. We shall then define a notion of *subformula* and use it to define a special class of *propositional formulas*. These latter help us to define the complex relation terms.

[6]For simplicity, we shall not allow definite descriptions with free variables in our language. (This is explained in footnote 7.) So, in what follows, we assume that the syntactic notion of a free variable occurring in a formula or term is also simultaneously defined.

> **Simple Terms**. Any constant or variable of type t is a (simple) term of type t.
>
> **Exemplification formulas**.
>
> > Where $\tau_1, \ldots, \tau_n$ (for $n \geq 0$), are terms of type $t_1, \ldots, t_n$, respectively, and Π is a term of type $\langle t_1, \ldots, t_n \rangle$, then the expression $\Pi\tau_1 \ldots \tau_n$ is an exemplification formula.
>
> When $n \geq 1$, we read $\Pi\tau_1 \ldots \tau_n$ as "$\tau_1, \ldots, \tau_n$ *exemplify* Π", and when $n = 0$, we read Π as "Π is true". Truth is the 0-place case of exemplification.
>
> **Encoding formulas**. There is one kind of *encoding* formula:
>
> > Where τ is any term of type t and Π is a term of type $\langle t \rangle$, then the expression $\tau\Pi$ is an encoding formula.
>
> We read $\tau\Pi$ as: τ encodes Π.
>
> **Complex formulas**.
>
> > Where φ, ψ are any formulas and α is any variable, then $\neg\varphi$ ('it is not the case that φ'), $\varphi \rightarrow \psi$ ('if φ, then ψ') and $\forall\alpha\varphi$ ('every α is such that φ') are complex formulas.
>
> We henceforth employ formulas of the form $\varphi \,\&\, \psi$, $\varphi \vee \psi$, and $\varphi \equiv \psi$, as these can be defined in terms of our complex formulas.
>
> **Subformulas**. We define *is a subformula of* φ as follows:
>
> 1. φ is a subformula of φ.
> 2. If $\neg\psi$ is a subformula of φ, then ψ is a subformula of φ.
> 3. If $\psi \rightarrow \chi$ is a subformula of φ, then ψ and χ are subformulas of φ.
> 4. If $\forall\alpha\psi$ is a subformula of φ, then ψ is a subformula of φ.
> 5. Nothing else is a subformula of φ.

We say that ψ is a *proper* subformula of φ just in case ψ is a subformula of φ but not identical to φ.

Propositional Formulas. φ is a *propositional formula* iff φ has no encoding subformulas.

Complex terms. There are two kinds of complex terms: (1) definite descriptions, and (2) complex relational terms.

1. Definite descriptions. Where α is any variable of type $t \neq \langle\,\rangle$ and φ is any formula in which α is the *only* variable that occurs free, then $\iota\alpha\varphi$ ("the α such that φ") is a complex term having type t.[7]

2. Complex relational terms. Where φ is any *propositional* formula, then (a) if $\alpha_1, \ldots, \alpha_n$ $(n \geq 1)$ are variables of type $t_1, \ldots, t_n$, respectively, then $[\lambda\alpha_1 \ldots \alpha_n\, \varphi]$ ("being an $\alpha_1, \ldots, \alpha_n$ such that φ") is a complex relation term having type $\langle t_1, \ldots, t_n\rangle$,[8] and (b) φ itself is a complex relation term having type $\langle\,\rangle$.

Although the foregoing defines the language $\mathcal{L}$ of our logical framework in complete generality, we shall frequently, in what follows, work with only a *fragment* of this language. For example, we often work with abstract objects denoted by expressions limited to the following types: i, $\langle\,\rangle$, $\langle i\rangle$, $\langle i, i\rangle$, $\langle\langle i\rangle\rangle$, and $\langle\langle i, i\rangle\rangle$. (In the Appendix, we define an explicit fragment by defining the bounded language $\mathcal{L}_{n,m}$, that includes these and the other types needed in what follows.) Thus, we'll be using the following specific variables:

[7]This rules out two kinds of descriptions. First it rules out a description such as $\imath p(p \,\&\, \neg p)$. In this example, the description operator would bind a variable of type $\langle\,\rangle$. We won't need descriptions of this type in our analysis of mathematics.

Second, the clause rules out open descriptions; we employ only *closed* descriptions in the language. There are three reasons for this. First, open descriptions won't be needed in the applications of the theory. Second, if descriptions with free variables don't occur in the language, then we need not worry about λ-expressions like those involved in the Kirchner Paradox (Kirchner 2017), where the λ binds a variable that occurs free in an encoding formula buried within a definite description that appears somewhere in the matrix of the λ-expression. Third, we can simplify the proof that our axioms are consistent.

[8]For purposes of this paper, we won't need terms of the form $[\lambda\, \varphi]$, where the λ doesn't bind any variables. But in other applications of object theory, these terms are allowed.

- $x, y, z, \ldots$ are variables of type i and so range over individuals
- $p, q, r, \ldots$ are variables of type $\langle\,\rangle$, and so range over propositions
- $F, G, H, \ldots$ are variables of type $\langle i\rangle$, and so range over first-level properties of individuals,
- $R, S, \ldots$ are variables of type $\langle i, i\rangle$, and so range over first-level relations among individuals
- $\mathcal{F}, \mathcal{G}, \mathcal{H}, \ldots$ are variables of type $\langle\langle i\rangle\rangle$, and so range over properties of properties of individuals
- $\mathcal{R}, \mathcal{S}, \ldots$ are variables of type $\langle\langle i, i\rangle\rangle$, and so range over properties of binary relations among individuals

Notice, here, that we've now used the symbol 'S' in two ways: earlier in the paper we used S as a constant to denote the property *being a set* (and thus an expression of type $\langle i\rangle$) and in the above list of variables, we've used S as a variable ranging over first-level relations (and thus an expression of type $\langle i, i\rangle$). The context will always make it clear which of these is intended.

To elucidate the above definitions, a series of observations is in order. First, where S is the constant of type $\langle i\rangle$ for *being a set* and $\emptyset$ is a constant of type i, then $S\emptyset$ is an exemplification formula ('$\emptyset$ exemplifies being a set'). Similarly, where $\in$ is a constant of type $\langle i, i\rangle$ and $\mathcal{R}$ is a variable of type $\langle\langle i, i\rangle\rangle$, then $\mathcal{R}\in$ is also an exemplification formula. If y is a variable of type i, $\emptyset$ is a constant of type i and $\in$ a constant of type $\langle i, i\rangle$, then $y \in \emptyset$ is also an exemplification formula, given our convention of rewriting relational predications using infix notation for the membership relation. Formulas like these made an appearance when we were setting up the illustrative examples of theoretical identifications in the previous subsection.

Second, it may be of interest to review several of the examples of encoding formulas already encountered, such as xF, $_F\mathcal{F}$, and $_R\mathcal{R}$. The formula xF is well-formed because x has type i and F has type $\langle i\rangle$. The formula $_F\mathcal{F}$ is well-formed because F has type $\langle i\rangle$ and $\mathcal{F}$ has type $\langle\langle i\rangle\rangle$. By the convention introduced above, we write the argument term of an encoding formula in a slightly smaller font, to better distinguish it from the relation term. The same goes for the encoding formula $_R\mathcal{R}$, which again is well-formed because R has type $\langle i, i\rangle$ and $\mathcal{R}$ has type $\langle\langle i, i\rangle\rangle$.

Third, note that the propositional formulas are those formulas which are built up out of exemplification formulas and the sentence-forming operations of negation, conditionalization, and quantification described in the definition of complex formulas. Consequently encoding formulas can only make an appearance inside a propositional formula φ if they are buried in a term within some propositional subformula of φ. For example, the formula $Rx\imath y(yG)$ ('x and $\imath y(yG)$ exemplify the relation R') and the formula $[\lambda x Rx\imath y(yG)]z$ ('z exemplifies the property of being an x that bears R to $\imath y(yG)$') are well-formed propositional formulas since they have no encoding subformulas. By contrast, the formula $\forall F(xF \rightarrow Fx)$ is not propositional, since it has xF as an encoding subformula.

Fourth, it should be helpful to know that we've introduced the restriction that banishes encoding subformulas from λ-expressions to avoid a Russell-style paradox. This paradox has been discussed in a variety of other publications (Zalta 1983, 1988, and more recently, Bueno, Menzel, & Zalta 2014), but for now, it suffices to note that if such expressions as $[\lambda x \exists F(xF \& \neg Fx)]$ were well-formed, then a contradiction would be derivable from the assertion that there is an abstract individual that encodes such a property. (An abstract individual that encodes $[\lambda x \exists F(xF \& \neg Fx)]$ would exemplify this property iff it does not.)

Fifth, since the variables $p, q, \ldots$ are terms of type $\langle\,\rangle$, they are also formulas, by the definition of exemplification formulas. Thus, by the second clause of the definition of complex terms, we can form λ-expressions such as $[\lambda x\, p]$. These denote a properties of individuals, i.e., a property with type $\langle i \rangle$, and we read $[\lambda x\, p]$ as *being such that* p, where p may denote any proposition.

Finally, by the second clause of the definition of complex terms, we shall be able to formulate λ-expressions such as $[\lambda y\, \varphi]$, $[\lambda F\, \varphi]$, and $[\lambda R\, \varphi]$, when φ is propositional. These will denote, respectively, a property of individuals, a property of first-level properties, and a property of first-level relations. Note that the variable bound by the λ need not be free in φ. As we shall see, the resulting expressions behave as expected. For example, it is axiomatic that in the case where the variable y is not free in φ, an individual x exemplifies $[\lambda y\, \varphi]$ iff (the proposition denoted by) φ is true.

Given these observations about the language of our logical framework, we conclude this section by:

(1) defining the property *being ordinary* to the negation of $A!$,

(2) distinguishing between abstract and ordinary objects of every type by stating their identity conditions, and

(3) defining the conditions under which a proposition p is *true in* an abstract individual x.

Concerning (1). We previously mentioned that where t is any type, then '$A!$' is a distinguished predicate of type $\langle t \rangle$. The symbol $A!$ is a 'typically ambiguous' primitive that denotes a property exemplified by the objects of type t that are *abstract*. And where t is any type and α is a variable of type t, we say that the property *being ordinary* ('$O!$') is *being an* α *such that* α *fails to exemplify being abstract*:

$$O! \ =_{df} \ [\lambda \alpha\, \neg A!\alpha]$$

Thus, the typically ambiguous predicate $O!$ is an expression of type $\langle t \rangle$, for any type t. The predicates $A!$ and $O!$ consequently partition the domain of each type t into the abstract and ordinary objects of type t.

Note that identity is not among the primitives of our logical framework. Identity can instead be defined. Although the definitions in full generality are complex, they are easy to grasp. If x and y are any abstract objects of type t, where t is any type, then x and y are identical whenever x and y encode the same properties having type $\langle t \rangle$. If x and y are any ordinary objects, then we define their identity by cases: (a) ordinary individuals x and y are identical whenever they exemplify the same properties; (b) ordinary properties F and G with type $\langle t \rangle$, where t is any type, are identical just in case they are encoded by the same objects. Identity for ordinary objects of the remaining types are defined in terms of property identity: (c) ordinary propositions p and q of type $\langle\,\rangle$ are identical just in case the properties *being an individual such that* p and *being an individual such that* q are identical; and (d) ordinary relations F and G of type $\langle t_1, \ldots, t_n \rangle$, where $t_1, \ldots, t_n$ are any types, are identical just in case every way of projecting F and G onto any $n-1$ objects of the right type yields identical properties.

Readers who are satisfied with these may skip ahead to the discussion of (3), but we now give a completely general definition by cases. For individuals, we say:

> Objects of type i are identical whenever they are either both ordinary objects of type i and exemplify the same properties or both abstract objects of type i and encode the same properties.

I.e., where x and y are both distinct variables of type i, and $O!$, $A!$, and F have type $\langle i\rangle$, we capture this definition as follows:

$$x\!=\!y \;=_{df}\; (O!x \;\&\; O!y \;\&\; \forall F(Fx \equiv Fy)) \vee (A!x \;\&\; A!x \;\&\; \forall F(xF \equiv yF))$$

Now where F and G are both of type $\langle t\rangle$ (i.e., they have the type of a property), for any type t, we define:

> F and G are *identical* if and only if either (a) F and G are both ordinary properties and are encoded by the same objects of type t or (b) F and G are both abstract and necessarily encode the same properties having type $\langle\langle t\rangle\rangle$

I.e., where $O!$, $A!$, and $\mathcal{H}$ are all of type $\langle\langle t\rangle\rangle$:

$$F\!=\!G \;=_{df}\; (O!F \;\&\; O!G \;\&\; \forall x^t(xF \equiv xG)) \vee (A!F \;\&\; A!G \;\&\; \forall\mathcal{H}(F\mathcal{H} \equiv G\mathcal{H}))$$

And where p and q are both of type $\langle\,\rangle$ (i.e., they have the type of a proposition), we define:

> p and q are *identical* if and only if either (a) p and q are both ordinary propositions and the propositional properties $[\lambda x\, p]$ and $[\lambda x\, q]$ are identical or (b) p and q are both abstract and encode the same properties having type $\langle\langle\,\rangle\rangle$

I.e., where x is a variable of type i and $O!$, $A!$, and $\mathcal{H}$ have type $\langle p\rangle$:

$$p\!=\!q \;=_{df}\; (O!p \;\&\; O!q \;\&\; [\lambda x\; p]\!=\![\lambda x\; q]) \vee (A!p \;\&\; A!q \;\&\; \forall\mathcal{H}(p\mathcal{H} \equiv q\mathcal{H}))$$

Finally, if F and G are of relational type $\langle t_1,\ldots t_n\rangle$, where $t_1,\ldots,t_n$ are any types, then F is identical to G 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 and encode the same properties with type $\langle\langle t_1,\ldots,t_n\rangle\rangle$. We leave the formal representation to a footnote.[9]

[9]Formally:

$$F^{\langle t_1,\ldots t_n\rangle}\!=\!G^{\langle t_1,\ldots t_n\rangle} \;=_{df}\; \text{(where } n>1)$$

$$\begin{aligned}&O!F \;\&\; O!G \;\&\; \forall y^{t_2}\ldots\forall y^{t_n}([\lambda x^{t_1}\; Fx^{t_1}y^{t_2}\ldots y^{t_n}]\!=\![\lambda x^{t_1}\; Gx^{t_1}y^{t_2}\ldots y^{t_n}]) \;\&\\ &\forall y^{t_1}\forall y^{t_3}\ldots\forall y^{t_n}([\lambda x^{t_2}\; Fy^{t_1}x^{t_2}y^{t_3}\ldots y^{t_n}]\!=\![\lambda x^{t_2}\; Gy^{t_1}x^{t_2}y^{t_3}\ldots y^{t_n}]) \;\&\ldots\&\\ &\qquad\forall y^{t_1}\ldots\forall y^{t_{n-1}}([\lambda x^{t_n}\; Fy^{t_1}\ldots y^{t_{n-1}}x^{t_n}]\!=\![\lambda x^{t_n}\; Gy^{t_1}\ldots y^{t_{n-1}}x^{t_n}]) \quad\vee\\ &A!F \;\&\; A!G \;\&\; \forall\mathcal{H}(F\mathcal{H} \equiv G\mathcal{H})\end{aligned}$$

Concerning (3). Using the notion of identity just defined, we may define two more notions that are needed to see how the framework parses the three theoretical identifications in the illustrative examples of Section 2.1. First, we define a *situation* to be any individual x such that every property x encodes is a property of the form $[\lambda y\, p]$, for some proposition p. Formally, where x and y are variables of type i and F is a variable of type $\langle i\rangle$, then:

$$\mathit{Situation}(x) \;=_{df}\; \forall F(xF \rightarrow \exists p(F\!=\![\lambda y\, p]))$$

Then where s is any situation, we say p *is true in* s, written $s \models p$, iff s encodes the property of *being-such-that-p*:

$$s \models p \;=_{df}\; s[\lambda y\, p]$$

Note that since $s \models p$ is defined in terms of the encoding formula $s[\lambda y\, p]$, it may not appear as a subformula in a propositional formula.

With this definition, the following three illustrative examples of theoretical identifications described earlier may be formally parsed in our logical framework, if given $\textsc{zf}$ and $\emptyset_{\mathrm{ZF}}$ as terms of type i, S_{ZF} as a term of type $\langle i\rangle$, and $\in_{\mathrm{ZF}}$ as a term of type $\langle i,i\rangle$:

$$\emptyset_{\mathrm{ZF}} = \iota x(A!x \;\&\; \forall F(xF \equiv \textsc{zf} \models F\emptyset_{\mathrm{ZF}}))$$

$$S_{\mathrm{ZF}} = \iota F(A!F \;\&\; \forall\mathcal{F}(F\mathcal{F} \equiv \textsc{zf} \models \mathcal{F}S_{\mathrm{ZF}}))$$

$$\in_{\mathrm{ZF}} \;=\; \iota R(A!R \;\&\; \forall\mathcal{R}(R\mathcal{R} \equiv \textsc{zf} \models \mathcal{R}\in_{\mathrm{ZF}}))$$

In Section 3, we discuss the axioms required to guarantee that each definite descriptions in the above denotes an abstract object of the appropriate type, and then in Section 4, we discuss how formulas having the form $\textsc{zf} \models F\emptyset_{\mathrm{ZF}}$, $\textsc{zf} \models \mathcal{F}S_{\mathrm{ZF}}$, and $\textsc{zf} \models \mathcal{R}\in_{\mathrm{ZF}}$, as well as the theoretical identifications above, become theorems of our logicist account of mathematics.

3 The Axioms for the Logical Framework

We can reason using the preceding logical framework by adopting the following groups of principles and rules:

1. The classical axioms and rules of predicate logic, as they are formulated for relational type theory. These are modified only to accom-

modate the (negative) free logic of definite descriptions.[10] (Thus, a definite definition $\iota\alpha\varphi$ can be instantiated into universal claims only when it is known, by proof or by hypothesis, that $\exists\beta(\beta\!=\!\iota\alpha\varphi)$, i.e., that the description is logically proper.[11])

2. The axiom governing the defined notion of identity.

3. The axioms governing the two kinds of complex terms: definite descriptions and the λ-expressions.

4. The axioms governing the primitive predicate $A!$ and governing encoding predications.

We shall not review the classical axioms and rules of relational type theory. Suffice it to say that all of the usual axioms and rules of propositional logic are included, and that the classical quantifier axioms and rules (suitably modified to accommodate the free logic of definite descriptions) apply to all formulas with quantifiers over variables of any type. In addition to these axioms, we assume only the primitive rules of Modus Ponens and Generalization, and the usual rules that are derivable from this basis.

It is important to emphasize here, however, that our framework and its application do not semantically presuppose anything more than general Henkin models. The first-level property variables $F, G, \ldots$ need not range over the full power set of the domain over which the individual variables $x, y, \ldots$ range. And, in general, our model in the Appendix shows that the domain of properties having type $\langle t\rangle$ is not the power set of the domain of

[10] We note that the logic of complex λ-expressions is classical, even though such terms might contain non-denoting descriptions. For even if they contain non-denoting descriptions, they will be assigned a denotation, i.e., an element of the appropriate domain. That's because their denotations are assigned on the basis of the satisfaction conditions of their matrix, which are well-defined even when the descriptions these matrices may contain denote nothing. See also our definition of Denotation in the Appendix.

[11] We shall assume familiarity with the following facts about free logic. First, the classical quantifier axiom for universal instantiation is modified so that terms τ can only be instantiated into a universal claim if one knows that $\exists\beta(\beta = \tau)$. Second, for every term τ other than a description, it is axiomatic that $\exists\beta(\beta = \tau)$. Third, and finally, for definite descriptions of the form $\iota\alpha\varphi$ it is an axiom that: $\psi^{\iota\alpha\varphi}_{\alpha} \rightarrow \exists\beta(\beta = \iota\alpha\varphi)$, where ψ is any atomic exemplification or encoding formula in which α occurs as one of the arguments, β doesn't occur free in φ, and $\psi^{\iota\alpha\varphi}_{\alpha}$ is the result of substituting $\iota\alpha\varphi$ for all the free occurrences of α in φ. This simply captures the principle underlying negative free logic that any atomic formula containing a non-denoting term is false.

objects of type t.[12] Nevertheless, in the model described in the Appendix, the axioms discussed below are all true.

In the remainder of this section, then, we describe the axioms that govern our defined notion of identity (Section 3.1), that govern the complex terms (Section 3.2), and that govern ordinary and abstract objects of any type t (Section 3.3).

3.1 Substitution of Identicals

As to identity, note first that given our definition of identity in the previous section, one can derive formulas of the form $\alpha = \alpha$ from the classical axioms and rules of our logic. The derivation is straightforward and will not be provided here. The following axiom ensures that when objects are identical, anything true about the one is true about the other, and vice versa:

> $\alpha\!=\!\beta \rightarrow [\varphi \equiv \varphi']$, where α, β are distinct variables of the same type and φ' is the result of replacing zero or more free occurrences of α in φ with occurrences of β.

Thus, our define notion of identity behaves classically, for every logical type.

3.2 Axioms Governing the Complex Terms

3.2.1 Definite Descriptions

The principle governing definite descriptions is simply this:

> AXIOM 1 (**Description Axiom**).
> $\beta\!=\!\iota\alpha\varphi \equiv \forall\alpha(\varphi \equiv \alpha\!=\!\beta)$, provided β is substitutable for α in φ.

This asserts: β is the α such that φ if and only if β is uniquely φ. As a simple example, let α, β be the type i variables x, y, respectively, let Q be a type $\langle i\rangle$ constant, and let φ be the exemplification formula Qx. Then the following is an instance of the Description Axiom:

$$y\!=\!\iota xQx \equiv \forall x(Qx \equiv x\!=\!y)$$

[12] In general, the domain $\mathbf{D}_t$ of type t is the union of the ordinary objects ot type t and the abstract objects of type t. So, the properties in $\mathbf{D}_{\langle t\rangle}$ includes all the ordinary properties with type $\langle t\rangle$ and the abstract properties with type $\langle t\rangle$. It will be seen, upon inspection, that this is *not* the power set of $\mathbf{D}_t$.

This asserts: y is identical to the x such that xQ if and only if y is the unique individual that exemplifies Q. Although we shall not take the time to prove it here, the classical Russell axiom for descriptions is now derivable.[13]

3.2.2 Principle Governing Relations

We employ the standard axiom of β-Conversion for relations denoted by λ-expressions in which the λ binds one or more variables. To state β-Conversion, let $\varphi^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$ be the result of substituting the term τ_i for every free occurrence of α_i in φ $(1 \leq i \leq n)$:

> AXIOM 2: $[\lambda\alpha_1 \ldots \alpha_n\, \varphi]\beta_1 \ldots \beta_n \equiv \varphi^{\beta_1,\ldots,\beta_n}_{\alpha_1,\ldots,\alpha_n}$, provided β_i is of the same type as, and substitutable for, α_i in φ.

This is just the familiar β-Conversion for n-place relations $(n \geq 1)$, and it governs the meaning of the term forming-operator $\lambda\alpha_1 \ldots \alpha_n$ ('being $\alpha_1, \ldots, \alpha_n$ such that'). Note that AXIOM 2 has instances for any appropriate types. Where F, G are properties that have type $\langle i \rangle$, and S, R are relations that have type $\langle i, i \rangle$, the following are consequences of AXIOM 2, by universal generalization:

(E) $\forall G([\lambda \mathit{F}\, \varphi]\mathit{G} \equiv \varphi^{G}_{F})$, provided G is substitutable for F in φ

(F) $\forall S([\lambda \mathit{R}\, \varphi]\mathit{S} \equiv \varphi^{S}_{R})$, provided S is substitutable for R in φ

As a specific example of (E), suppose 'S_1' denotes the property of being a set, so that 'S_1' is being used as a constant of type $\langle i \rangle$. (We haven't yet said what this property is, and in fact, we haven't assumed, and won't assume, that there is a single, unique property of being a set. But with this proviso, it shouldn't be misleading if we continue to develop our example.) Then the following is a consequence of (E), in which we've instantiated the universal quantifier $\forall G$ in (E) to the specific property S_1:

$$[\lambda \mathit{F}\, \neg\exists y(\mathit{F}y \,\&\, y \in \emptyset)]\mathit{S}_1 \equiv \neg\exists y(\mathit{S}_1 y \,\&\, y \in \emptyset)$$

And if we let $\in$ be a membership relation, then the following is a consequence of (F), in which the we've instantiated the universal quantifier $\forall R$ of AXIOM 2 to the specific relation $\in$:

[13] The proof depends on the axiom required for the free logic of descriptions noted in footnote 11.

$$[\lambda \mathit{R}\, \neg\exists y(\mathit{F}y \,\&\, y\mathit{R}\emptyset)]{\in} \equiv \neg\exists y(\mathit{F}y \,\&\, y \in \emptyset)$$

Thus, instances of β-Conversion simply require that the denotation of the λ-expression to be a relation whose exemplification extension is equivalent to the λ-expression's matrix. We'll also assume that a principle of α-Conversion, which asserts an identity between alphabetic variants, governs all our λ-expressions; but for simplicity, we omit this axiom.

3.3 Principles Governing Encoding

We turn finally to the axioms governing our primitive predicate $A!$ in both exemplification and encoding predications.

3.3.1 What Is Abstract

First, we introduce the axioms that assert the existence of abstract objects of every type. Where α is a variable of type t, F is a variable of type $\langle t \rangle$, and $A!$ is a predicate of type $\langle t \rangle$, we assert:

> AXIOM 3: $\exists\alpha(A!\alpha \,\&\, \forall F(\alpha F \equiv \varphi))$, where φ has no free αs.

Here are three examples. In the first, x is a variable of type i, while $A!$ and F are of type $\langle i \rangle$. In the second, F is a variable of type $\langle i \rangle$, while $A!$ and $\mathcal{F}$ are of type $\langle\langle i \rangle\rangle$. In the third, R is a variable of type $\langle i, i \rangle$, while $A!$ and $\mathcal{R}$ are of type $\langle\langle i, i \rangle\rangle$:

$\exists x(A!x \,\&\, \forall F(xF \equiv \varphi))$, where φ has no free Fs

$\exists \mathit{F}(A!\mathit{F} \,\&\, \forall \mathcal{F}(\mathit{F}\mathcal{F} \equiv \varphi))$, where φ has no free Fs

$\exists \mathit{R}(A!\mathit{R} \,\&\, \forall \mathcal{R}(\mathit{R}\mathcal{R} \equiv \varphi))$, where φ has no free Rs

The first asserts that there exists an abstract individual that encodes all and only the properties of individuals that satisfy φ. The second asserts that there exists an abstract property of individuals that encodes all and only the properties of properties of individuals that satisfy φ. The third asserts that there exists an abstract relation among individuals that encodes exactly properties of relations among individuals that satisfy φ.

Notice that from any instance of the above, we can derive the existence of a unique such object by an appeal to the definition of abstract object identity. Consider the second example above. There couldn't be two distinct abstract properties of individuals that encode exactly the properties of properties satisfying φ, since distinct abstract properties, by definition, have to differ by one of their encoded properties.

Thus the above principles guarantee that the definite descriptions used in our three illustrative examples of theoretical identifications outlined in Section 2.1 are logically proper or well-defined (i.e., have denotations), since they are constructed in terms of formulas φ that have no free xs, Fs or Rs, respectively. In other words, the following are theorems:

$$\exists y(y = \iota x(A!x \,\&\, \forall F(xF \equiv \varphi)))$$

$$\exists G(G = \iota F(A!F \,\&\, \forall \mathcal{F}(F\mathcal{F} \equiv \varphi)))$$

$$\exists S(S = \iota R(A!R \,\&\, \forall \mathcal{R}(R\mathcal{R} \equiv \varphi)))$$

In what follows, we call such descriptions *canonical* since for any formula φ (excluding only those with an inappropriate variable), the descriptions are guaranteed to have a denotation.

Moreover, the following abstraction principles are derivable as theorems that govern canonical descriptions:

$\iota x(A!x \,\&\, \forall F(xF \equiv \varphi))G \equiv \varphi'$, where G is free for F in φ and φ' is the result of substituting G for F everywhere in φ.

$\iota F(A!F \,\&\, \forall \mathcal{F}(F\mathcal{F} \equiv \varphi))\mathcal{G} \equiv \varphi'$, where $\mathcal{G}$ is free for $\mathcal{F}$ in φ and φ' is the result of substituting $\mathcal{G}$ for $\mathcal{F}$ everywhere in φ.

$\iota R(A!R \,\&\, \forall \mathcal{R}(R\mathcal{R} \equiv \varphi))\mathcal{S} \equiv \varphi'$, where $\mathcal{S}$ is free for $\mathcal{R}$ in φ and φ' is the result of substituting $\mathcal{S}$ for $\mathcal{R}$ everywhere in φ.

The first theorem asserts that the abstract individual encoding just the properties such that φ encodes a property G if and only if G is such that φ. The second theorem asserts that the abstract property of individuals encoding just the properties such that φ encodes a property $\mathcal{G}$ if and only if $\mathcal{G}$ is such that φ. The third theorem asserts that the abstract relation encoding just the properties such that φ encodes a property $\mathcal{S}$ if and only if $\mathcal{S}$ is such that φ.

To state the axiom, we introduce a distinguished abstract object by way of a canonical description. Let us say that *the null object* of type t, written $\boldsymbol{a}^t_\emptyset$, is that abstract object of type t that encodes all and only non-self-identical properties.

$$\boldsymbol{a}^t_\emptyset =_{df} \iota\alpha(A!\alpha \,\&\, \forall F(\alpha F \equiv F \neq F))$$

$\boldsymbol{a}^t_\emptyset$ is indeed a null object since it provably encodes no properties. We may now assert, for any object α of type t, that if α is not identical to the the null object, then α is abstract if and only if α encodes a property. Where α is of type t, and F and $A!$ are of type $\langle t \rangle$, this axiom may be formalized as follows:

AXIOM 4: $\alpha \neq \boldsymbol{a}^t_\emptyset \rightarrow (A!\alpha \equiv \exists F\alpha F)$

Note, as a simple corollary, that $\exists F\alpha F \rightarrow A!\alpha$, for if α encodes a property, it is not the null object, and so the right-to-left direction of the consequent of AXIOM 4 implies it is abstract.

Suppose, for example, that our language included the name s for the ordinary individual Socrates and that, as a premise, we suppose $\neg A!s$. Then it follows that Socrates is not the null object and, moreover, that Socrates fails to encode properties.

3.3.2 What Isn't Abstract

Our remaining axioms tell us about what isn't abstract. From AXIOM 4, we know that other than the null object, something is abstract entities if and only if it encodes properties; abstract objects objectify, at a lower level, higher-level patterns of properties already present in exemplification logic; they objectify the properties that satisfy higher-level conditions on properties. Basically, any expression that denotes something that doesn't objectify patterns by encoding properties are, intuitively, not abstract.

So, the first point to notice is that abstract relations, of any type, simply (a) encode properties and (b) exemplify properties of relations and relations among relations; however, *nothing* exemplifies them. So, if a relation is exemplified, it fails to be abstract. Where F is a variable of type $\langle t_1, \ldots, t_n \rangle$, $A!$ has type $\langle\langle t_1, \ldots, t_n \rangle\rangle$, and $\alpha_1, \ldots \alpha_n$ are distinct variables of type $t_1, \ldots, t_n$, respectively, then for $n \geq 0$, it is axiomatic that:

AXIOM 5: $\exists\alpha_1 \ldots \exists\alpha_n F\alpha_1 \ldots \alpha_n \rightarrow \neg A!F$

Note that in the case of the empty type $\langle\,\rangle$, this axiom implies that true propositions are not abstract (i.e., $p \rightarrow \neg A!p$), and hence that abstract propositions are false, i.e., that $A!p \rightarrow \neg p$.

AXIOM 5 also tells us that β-Conversion *never* applies to abstract properties and abstract relations, since nothing ever exemplifies them.

Intuitively, a property like $[\lambda x\, \varphi]$ is something that is exemplifiable by all and only the things satisfying φ, where φ is an exemplification pattern. But abstract properties and relations arise by comprehension, i.e., by what they encode, not by what exemplifies them. So λ-constructors build things that can be exemplified, and entities defined by what they encode aren't things that can be exemplified, and hence they are not abstract.[14] Thus we assert:

$$\text{AXIOM 6: } \neg A![\lambda\nu_1 \ldots \nu_n\, \varphi] \qquad (n \geq 1)$$

We leave the formulation of examples of AXIOM 6 to the reader.

Finally, in the special case where F is a variable of type $\langle t_1, \ldots, t_n\rangle$, and $\alpha_1, \ldots, \alpha_n$ are distinct variables of type $t_1, \ldots, t_n$, respectively, and $A!$ has type $\langle\langle t_1, \ldots, t_n\rangle\rangle$, then we also assert that η-Conversion holds for elementary λ-expressions in which the 'head' relation is not an abstract relation:

$$\text{AXIOM 7: } \neg A!F \rightarrow ([\lambda\alpha_1 \ldots \alpha_n\, F\alpha_1 \ldots \alpha_n] = F) \qquad (n \geq 1)$$

4 Application to Mathematics

To develop our logicist account of mathematics, we note first that by 'mathematics' we shall be focusing on *theoretical* as opposed to *natural* mathematics. *Natural* mathematics consists of the ordinary, pretheoretic claims that seem to be about mathematical objects, such as the following:

- The Triangle has 3 sides.
- The number of planets is eight.
- There are more individuals in the class of insects than in the class of humans.
- Lines a and b have the same direction.
- Figures a and b have the same shape.

[14]Later, we will introduce special λ-expressions indexed to a theory that denote abstract entities: while they capture exemplification patterns *relative to a theory*, they are nevertheless defined as entities by what they encode. So the semantics of these really are tied to object comprehension and so β-Conversion doesn't apply to them.

Theoretical mathematics, on the other hand, involves claims that occur in the context of some mathematical theory, no matter whether the axioms have been made explicit or left implicit and no matter whether the theory has been formalized or not. Examples of such claims are:

- The empty set is an element of the unit set of the empty set.
  [Said with reference to Zermelo-Fraenkel set theory.]
- 2 is less than or equal to π.
  [Said with reference to real number theory.]

Though our framework can be applied to the analysis of both natural and theoretical mathematics, our present focus is only on the latter. For a discussion of the former, and how the second-order modal version of the above framework can analyze such terms as 'The Triangle', 'the number of planets', 'the class of insects', etc., see Pelletier & Zalta 2000, Zalta 1999, and Anderson & Zalta 2004.[15]

We shall assume, in what follows, that to produce a logicist account of theoretical mathematics, we have to show that arbitrary mathematical theories can be *reduced* to logic plus analytic truths. If this is right, then our argument divides into two parts: (1) show that an arbitrary mathematical theory T can be reduced to the formal system described in Section 3, and (2) show that the formal system constitutes a logic supplemented with analytic truths. To achieve (1), at the very least, we have to (a) assign the terms and predicates of T denotations that are

[15]To give interested readers some hint, note that if we were to add modality to the present system, we could identify The Triangle ('Φ_T') as the abstract individual that encodes exactly the properties necessarily implied by being triangular (T):

$$\Phi_T \;=_{df}\; \iota x(A!x \,\&\, \forall F(xF \equiv \Box\forall y(Ty \rightarrow Fy)))$$

See Pelletier & Zalta 2000 for the details. We can identify the natural number of Gs ('$\#G$') as the abstract individual that encodes exactly those properties F that are in one-one correspondence with G on the ordinary objects ('$\approx_E$'):

$$\#G \;=_{df}\; \iota x(A!x \,\&\, \forall F(xF \equiv F \approx_E G))$$

See Zalta 1999 for the details, where it is shown how one can derive Hume's Principle ($\#F = \#G \equiv F \approx_E G$) from the above definition. Furthermore, we can identify the class of Gs ('ϵG') as the abstract individual that encodes exactly those Fs that are materially equivalent to G:

$$\epsilon G \;=_{df}\; \iota x(A!x \,\&\, \forall F(xF \equiv \forall y(Fy \equiv Gy)))$$

See Anderson & Zalta 2004 for the details, and for the proof of a consistent version of Basic Law V ($\epsilon F = \epsilon G \equiv \forall x(Fx \equiv Gx)$).

describable in our framework, and (b) assign the theorems of T a reading in our system, involving those denotations, on which they are true.

If (1) and (2) is sufficient to produce a logicist account of theoretical mathematics, then we can reap the epistemological benefits of logicism. In this section, we explain how the reduction of arbitrary mathematical theories is to be effected, and in the next section we argue that our formal framework is a logic. Though our framework is capable of analyzing mathematical theories of any finite order, recall that, for simplicity, we are targeting first- and second-order mathematical theories having only primitive constants and variables, and 1- and 2-place predicates (including λ-expressions, which are to be interpreted relationally), but without function terms, definite descriptions, or n-ary predicates for $n > 2$. Though our system is set to handle more complex kinds of theories, we need not be distracted here by the subtleties involved.

Our first step shall be to analyze what we pretheoretically judge to be (the content of) a mathematical theory as an abstract individual that encodes *propositions*. An abstract individual encodes a proposition p by encoding the propositional property of the form $[\lambda x\, p]$. We can identify a mathematical theory with an abstract object that encodes only propositions. This motivates the definition at the end of Section 2, where we stipulated that p *is true in* T ('$T \models p$') means that T encodes $[\lambda x\, p]$. Finally, we add constants of type i to our logical framework to denote what we pretheoretically judge to be mathematical theories: '$\textsc{zf}$' denotes Zermelo-Fraenkel set theory, '$\Re$' denotes real number theory, '$\textsc{pa}$' denotes Peano arithmetic, etc. It follows that statements of the form $\textsc{zf} \models p$, $\Re \models q$, $\textsc{pa} \models r$, etc., become well-formed.

Now the mechanism by which these statements become theorems of our framework is as follows. Consider an arbitrary mathematical theory T. We can import the theorems of T into our framework by appeal to the following Importation Principle (later we argue that the resulting claims are theory-relative analytic truths):

> **Importation Principle**. When φ is a closed theorem of T, then $T \models \varphi^*$ shall be an axiom, where φ^* is the result of indexing every occurrence of a term or predicate of T to T.

So, for example, given the Importation Principle, since it is the case that:

$$\vdash_{\mathrm{ZF}} \neg\exists y(Sy \,\&\, y \in \emptyset)$$

the following statement will be an axiom of our framework:

$$\textsc{zf} \models \neg\exists y(S_{\textsc{zf}}y \,\&\, y \in_{\textsc{zf}} \emptyset_{\textsc{zf}})$$

In what follows, we call the theory-indexed terms introduced into object theory by means of the Importation Principle T-*indexed* terms, or *indexed* terms for short, where T is the theory in question.

Now since we are assuming that mathematical theories have been formulated in a formal system that includes λ-expressions, we also assume the principle of λ-Conversion, which has the form $[\lambda\alpha\,\varphi]\kappa \equiv \varphi^{\kappa}_{\alpha}$, is part of the logic of T. Then our Importation Principle (by the logical closure of mathematical theories) yields the following special case, where $[\lambda\alpha\,\varphi^{\alpha}_{\kappa}]$ is the result of substituting the variable α for all the occurrences of κ in φ:

> **Closure Principle**. Let φ be a theorem of T, let κ be a constant term of type t occurring in φ, and let α be a variable of type t that doesn't occur free in φ, so that $[\lambda\alpha\,\varphi^{\alpha}_{\kappa}]$ is a 1-place relation term of T that has type $\langle t\rangle$. Then where φ^* is again the result of indexing to T all the terms of T and the theorems of T are imported into object theory, the following become axioms of object theory by importation:
>
> (a) $T \models ([\lambda\alpha\,(\varphi^{\alpha}_{\kappa})^*]_T\kappa_T \equiv \varphi^*)$
>
> (b) $T \models [\lambda\alpha\,(\varphi^{\alpha}_{\kappa})^*]_T\kappa_T$
>
> i.e., (a) it is true in T that: κ_T exemplifies $[\lambda\alpha\,(\varphi^{\alpha}_{\kappa})^*]_T$ if and only if φ^* is true in T, and (b) it is true in T that κ_T exemplifies $[\lambda\alpha\,(\varphi^{\alpha}_{\kappa})^*]_T$.

This shows that truth in a theory is closed under λ-Conversion. So when we import the theorems of T into object theory, the new axioms of object theory include those of the form $T \models [\lambda\alpha\,(\varphi^{\alpha}_{\kappa})^*]_T\kappa_T$.

Given (b), the Closure Principle ensures that the following are all axioms of object theory:

(G) $\textsc{zf} \models \neg\exists y(S_{\textsc{zf}}y \,\&\, y \in_{\textsc{zf}} \emptyset_{\textsc{zf}})$

(H) $\textsc{zf} \models [\lambda x\, \neg\exists y(S_{\textsc{zf}}y \,\&\, y \in_{\textsc{zf}} x)]_{\textsc{zf}}\emptyset_{\textsc{zf}}$

(I) $\textsc{zf} \models [\lambda F\, \neg\exists y(Fy \,\&\, y \in_{\textsc{zf}} \emptyset)]_{\textsc{zf}}S_{\textsc{zf}}$

(J) $\textsc{zf} \models [\lambda R\, \neg\exists y(S_{\textsc{zf}}y \,\&\, y\, R\, \emptyset_{\textsc{zf}})]_{\textsc{zf}}{\in_{\textsc{zf}}}$

Thus, not only does object theory become extended with axiom (G), i.e., in ZF, no set is a member of the nullset, but also with the axioms (H), (I), and (J), which assert that: (H) in ZF, the null set exemplifies the property *having no sets as members*, (I) in ZF, the property of being a set exemplifies the property of properties *being an F such that nothing exemplifying F is an element of the null set*, and (J) in ZF, the membership relation has the property of relations *being an R such that no set bears R to the null set*.

Now, at this point, we may theoretically identify the mathematical theory ZF using the following axiom:

$$\text{zf} = \iota x(A!x \,\&\, \forall F(xF \equiv \exists p(\text{zf} \models p \,\&\, F = [\lambda y\, p])))$$

In other words, zf is the abstract individual that encodes exactly the properties F such that there is a proposition p true in zf for which F is *being such that* p. In terms of our informal understanding of what it is for an abstract individual to encode a proposition, the above claim identifies ZF with that abstract individual that encodes exactly the propositions true in ZF. Similar identifications can be given for other mathematical theories. We now turn to the special axioms that identify denotations of the theoretical primitives (singular terms and predicates) of mathematical theories.

4.1 The Denotations of the Terms of T

Find a place to cite Leitgeb & Ladyman 2008, for the suggestiion that each theory T (they say 'structure') comes with its own identity relation.

We can now say what the constants and λ-expressions of T denote in terms of our background theory. Thus far, we've assumed T includes primitive constants, primitive 1-place predicates, primitive 2-place predicates, and λ-expressions. Let κ_T be the T-indexed version of the primitive constant κ of T, Π^1_T be the T-indexed version of the primitive 1-place predicate Π^1 of T, Π^2_T be the T-indexed version of the 2-place predicate Π^2 of T, and $[\lambda\alpha^t\, \varphi^*]_T$ be the T-indexed version of the 1-place λ-expression $[\lambda\alpha^t\,\varphi]$ of T. We then assert the following Reduction Axiom Schemata using canonical descriptions to identify the denotations of the new, indexed mathematical terms imported into object theory.

Reduction Axiom Schemata:

$$\kappa_T = \iota x(A!x \,\&\, \forall F(xF \equiv \ T \models F\kappa_T))$$

$$\Pi^1_T = \iota F(A!F \,\&\, \forall\mathcal{F}(F\mathcal{F} \equiv \ T \models \mathcal{F}\Pi^1_T))$$

$$\Pi^2_T = \iota R(A!R \,\&\, \forall\mathcal{R}(R\mathcal{R} \equiv \ T \models \mathcal{R}\Pi^2_T))$$

$$[\lambda\alpha\, \varphi^*]_T = \iota\beta(A!\beta \,\&\, \forall\Gamma(\beta\Gamma \equiv \ T \models \Gamma[\lambda\alpha\, \varphi^*]_T)),$$

where β is a variable of type $\langle t\rangle$, $A!$ has type $\langle\langle t\rangle\rangle$, and Γ is a variable of type $\langle\langle t\rangle\rangle$.

The latter asserts that the T-based property *being-an-*α^t*-such-that-*φ^* is identical to the abstract property that encodes all and only those properties of properties Γ exemplified (in T) by the T-based property *being-an-*α^t*-such-that-*φ^*. This abstract property has type $\langle t\rangle$.

Note that the above reduction axioms are *not* definitions, since the expressions on the left of the identity sign also appear on the right. But they are principles that are analytic, or so we will argue in the next section. Our next goal is to identify exactly which abstract objects are denoted by $\emptyset_{\text{ZF}}$, S_{ZF}, and $\in_{\text{ZF}}$. Then we'll work our way towards a derivation of some of the properties these abstract objects encode.

The following are instances of the reduction axioms that identify the primitive constants of ZF:

Instances of the Reduction Axiom Schemata:

$$\emptyset_{\text{ZF}} = \iota x(A!x \,\&\, \forall F(xF \equiv \ \text{zf} \models F\emptyset_{\text{ZF}}))$$

$$S_{\text{ZF}} = \iota F(A!F \,\&\, \forall\mathcal{F}(F\mathcal{F} \equiv \ \text{zf} \models \mathcal{F}S_{\text{ZF}}))$$

$$\in_{\text{ZF}} = \iota R(A!R \,\&\, \forall\mathcal{R}(R\mathcal{R} \equiv \ \text{zf} \models \mathcal{R}\in_{\text{ZF}}))$$

And the following are reduction axioms that identify some of the complex predicates of ZF. For the first example, we can reuse variables to obtain:

$$[\lambda x\, \neg\exists y(S_{\text{ZF}}y \,\&\, y \in_{\text{ZF}} x)]_{\text{ZF}} = \\ \iota F(A!F \,\&\, \forall\mathcal{F}(F\mathcal{F} \equiv \ \text{zf} \models \mathcal{F}[\lambda x\, \neg\exists y(S_{\text{ZF}}y \,\&\, y \in_{\text{ZF}} x)]_{\text{ZF}}))$$

This says that the ZF-property *having no sets as members* is identical to the abstract property that encodes all and only those second-level properties that are exemplified, in ZF, by the property of *having no sets as members*.

For the next identification, let F be a variable of type $\langle i\rangle$, let β be a variable of type $\langle\langle i\rangle\rangle$, let $A!$ have type $\langle\langle\langle i\rangle\rangle\rangle$, and let Γ be a variable of type $\langle\langle\langle i\rangle\rangle\rangle$. Then the following is also a reduction axiom:

$$[\lambda F\, \neg\exists y(Fy \,\&\, y \in_{\text{ZF}} \emptyset_{\text{ZF}})]_{\text{ZF}} = \\ \iota\beta(A!\beta \,\&\, \forall\Gamma(\beta\Gamma \equiv \ \text{zf} \models \Gamma[\lambda F\, \neg\exists y(Fy \,\&\, y \in_{\text{ZF}} \emptyset)]_{\text{ZF}}))$$

This says that the ZF-property of properties *being an F such that nothing exemplifying F is an element of the null set* is identical to the abstract second-level property (i.e., property of properties) that encodes all and only the third-level properties (i.e., properties of properties of properties) that are exemplified in ZF by the property *being an F such that nothing exemplifying F is an element of the null set.*

Finally, let R be a variable of type $\langle i, i\rangle$, let β be a variable of type $\langle\langle i, i\rangle\rangle$, let $A!$ have type $\langle\langle\langle i, i\rangle\rangle\rangle$, and let Γ be a variable of type $\langle\langle\langle i, i\rangle\rangle\rangle$. Then the following is also an instance of the Reduction Axiom schema:

$$[\lambda R\, \neg\exists y(S_{\text{ZF}}y \,\&\, y\, R\, \emptyset_{\text{ZF}})]_{\text{ZF}} = \\ \iota\beta(A!\beta \,\&\, \forall\Gamma(\beta\Gamma \equiv \text{ ZF} \models \Gamma[\lambda R\, \neg\exists y(S_{\text{ZF}}y \,\&\, y\, R\, \emptyset_{\text{ZF}})]_{\text{ZF}}))$$

As an exercise, the reader wish to formulate a gloss of this identification using the two previous examples as a guide.

Now let's focus just on the identification of the primitive constants and primitive predicates of a particular theory, so that we can see the consequences of the foregoing. The following Equivalence Theorem Schemata are immediate consequences of our Reduction Axiom Schemata, by the Abstraction Principles for abstract objects and substitution of identicals:[16]

Equivalence Theorem Schemata:

$$\forall G(\kappa_T G \equiv T \models G\kappa_T) \\ \forall\mathcal{G}(\Pi^1_T\mathcal{G} \;\equiv\; T \models\; \mathcal{G}\Pi^1_T) \\ \forall\mathcal{S}(\Pi^2_T\mathcal{S} \;\equiv\; T \models\; \mathcal{S}\Pi^2_T)$$

In other words, for any first-level property G, κ_T encodes G iff κ_T exemplifies G in T; for any second-level property of properties $\mathcal{G}$, the property Π^1_T encodes $\mathcal{G}$ iff Π^1_T exemplifies $\mathcal{G}$ in T, and for any second-level property of relations $\mathcal{S}$, the property Π^2_T encodes $\mathcal{S}$ iff Π^2_T exemplifies $\mathcal{S}$ in T.

Clearly, then, the following are instances of the Equivalence Theorems:

[16]Here, for example, is a proof of the second theorem. The Abstraction Principle for abstract properties has the following instance, where Π^1_T is any primitive 1-place predicate of T, and $\mathcal{G}$ is a free variable ranging over second-level properties of properties:

$$\iota F(A!F \,\&\, \forall\mathcal{F}(F\mathcal{F} \equiv T \models\; \mathcal{F}\Pi^1_T))\mathcal{G} \;\equiv\; T \models \mathcal{G}\Pi^1_T$$

But by the relevant Reduction Axiom, it follows by substitution of identicals that:

$$\Pi^1_T\mathcal{G} \;\equiv\; T \models \mathcal{G}\Pi^1_T$$

We can now universally generalize on the free variable $\mathcal{G}$ to reach our theorem.

$$\forall G(\emptyset_{\text{ZF}}G \;\equiv\; \text{ZF} \models G\emptyset_{\text{ZF}}) \\ \forall\mathcal{G}(S_{\text{ZF}}\mathcal{G} \;\equiv\; \text{ZF} \models\; \mathcal{G}S_{\text{ZF}}) \\ \forall\mathcal{S}(\in_{\text{ZF}}\mathcal{S} \;\equiv\; \text{ZF} \models\; \mathcal{S}\in_{\text{ZF}})$$

In other words, the empty set of ZF encodes exactly the properties G that it exemplifies in ZF; the ZF-property of being a set encodes exactly the second-level properties of properties that it exemplifies in ZF; and the membership relation of ZF encodes exactly the second-level properties of relations that it exemplifies in ZF.

Now the axioms resulting from the Importation Principle and specified by the Closure Principle become salient. For the properties that can be abstracted from those claims can be instantiated into the above universal claims. In particular, the properties that are referenced in (H), (I), and (J) above may be instantiated, respectively, into the above claims to yield the following theorems:

$$\emptyset_{\text{ZF}}[\lambda x\, \neg\exists y(S_{\text{ZF}}y \,\&\, y \in_{\text{ZF}} x)]_{\text{ZF}} \;\equiv\; \text{ZF} \models [\lambda x\, \neg\exists y(S_{\text{ZF}}y \,\&\, y \in_{\text{ZF}} x)]_{\text{ZF}}\emptyset_{\text{ZF}}$$

$$S_{\text{ZF}}[\lambda F\, \neg\exists y(Fy \,\&\, y \in_{\text{ZF}} \emptyset_{\text{ZF}})]_{\text{ZF}} \;\equiv\; \text{ZF} \models\; [\lambda F\, \neg\exists y(Fy \,\&\, y \in_{\text{ZF}} \emptyset_{\text{ZF}})]_{\text{ZF}}S_{\text{ZF}}$$

$$\in_{\text{ZF}}[\lambda R\, \neg\exists y(S_{\text{ZF}}y \,\&\, yR\emptyset_{\text{ZF}})]_{\text{ZF}} \;\equiv\; \text{ZF} \models\; [\lambda R\, \neg\exists y(S_{\text{ZF}}y \,\&\, yR\emptyset_{\text{ZF}})]_{\text{ZF}}\in_{\text{ZF}}$$

Each formal term in the biconditionals displayed above has been given a formal identification in our theory. Moreover, since the right-hand side of each of the above equivalences is a theorem resulting from the Importation and Closure Principles, we have a proof of the following facts about $\emptyset_{\text{ZF}}$, S_{ZF}, and $\in_{\text{ZF}}$:

$$\emptyset_{\text{ZF}}[\lambda x\, \neg\exists y(S_{\text{ZF}}y \,\&\, y \in_{\text{ZF}} x)]_{\text{ZF}} \\ S_{\text{ZF}}[\lambda F\, \neg\exists y(Fy \,\&\, y \in_{\text{ZF}} \emptyset_{\text{ZF}})]_{\text{ZF}} \\ \in_{\text{ZF}}[\lambda R\, \neg\exists y(S_{\text{ZF}}y \,\&\, yR\, \emptyset_{\text{ZF}})]_{\text{ZF}}$$

In other words, it is provable in our framework that the empty set of ZF encodes the ZF-property of having no sets as members; the ZF-property of being a set encodes the (second-level) ZF-property of being a property such that nothing exemplifying it is a member of the empty set; and the membership relation of ZF encodes the (second-level) ZF-property of being a relation that no set bears to the empty set.

4.2 Sentence Reduction: Truth Conditions of Mathematical Statements

So what reading is to be assigned to the theorems of mathematical theories when we consider those theorems in and of themselves, unprefixed by a theory operator? Platonist philosophers of mathematics believe that the unadorned claims of mathematics, such as '0 is a number', 'the empty set is an element of unit set of the empty set', 'two is less than π', etc. are simply true, while fictionalist philosophers argue that they are false. Our view is that this disagreement is explained by the fact that these claims are ambiguous, for there is an exemplification reading on which they are false and an encoding reading on which they are true.

Take a simple atomic formula, e.g., the statement that '0 is a number', when this is asserted as an axiom of Peano Arithmetic. It should be clear thus far that we take the claim that "In Peano Arithmetic, 0 is a number" to be true *simpliciter*, where in the first instance, this would be represented as:

$$\vdash_{\mathrm{PA}} N0$$

After importing the above into object theory, our analysis is:

$$\textsc{pa} \models N_{\mathrm{PA}} 0_{\mathrm{PA}}$$

But we can go a step further and give a true reading of the unprefixed "0 is a number". For given our Equivalence Theorems, it should be clear that the encoding formula, $0_{\mathrm{PA}} N_{\mathrm{PA}}$, is derivable as a theorem. We now have a precise and valid argument for why $0_{\mathrm{PA}} N_{\mathrm{PA}}$ is a true reading of "0 is a number". No such argument can be given for the exemplification reading, $N_{\mathrm{PA}} 0_{\mathrm{PA}}$, of the unadorned claim "0 is a number". In our framework, it is axiomatically false, by the contrapositive of AXIOM 6 and the fact that N_{PA} is an abstract property of individuals. We interpret this example to show that predications of the form 'x is F' in natural language are structurally ambiguous, and that in the case at hand, the encoding reading xF is provably true while the exemplification reading Fx is false.

Moreover, we take there to be a structural ambiguity in simple predications of natural language, for the unadorned claim "0 is a number" embodies not only a true atomic fact about a property that 0_{PA} encodes but also a true atomic fact about a property that N_{PA} encodes, namely, $[\lambda F\, F0_{\mathrm{PA}}]_{\mathrm{PA}}$. Once we import $\vdash_{\mathrm{PA}} [\lambda F\, F0]N$ so as to yield the axiom $\textsc{pa} \models [\lambda F\, F0_{\mathrm{PA}}]_{\mathrm{PA}} N_{\mathrm{PA}}$, the Equivalence Theorem guarantees that the encoding formula $N_{\mathrm{PA}}[\lambda F\, F0_{\mathrm{PA}}]_{\mathrm{PA}}$ is a theorem of our logical framework. Thus, we have another true reading of our unadorned mathematical claim. In a manner similar to the above, the exemplification reading, that N_{PA} exemplifies $[\lambda F\, F0_{\mathrm{PA}}]_{\mathrm{PA}}$, is axiomatically false, by AXIOM 6 and the fact that $[\lambda F\, F0_{\mathrm{PA}}]_{\mathrm{PA}}$ is an abstract property of properties.

To see how to generalize the procedure for assigning true encoding formulas as the truth conditions for unadorned mathematical theorems, let us return to the statement, "No set is an element of the empty set", said in the context of ZF. We've seen how the representation of the claim that this is a theorem of ZF gets imported as the following theorems of our logical framework (the first of which is an axiom by importation, the remainder of which are consequences):

$$\begin{aligned}
&\mathrm{ZF} \models \neg\exists y(S_{\mathrm{ZF}}y \,\&\, y \in_{\mathrm{ZF}} \emptyset_{\mathrm{ZF}})\\
&\mathrm{ZF} \models [\lambda x\, \neg\exists y(S_{\mathrm{ZF}}y \,\&\, y \in_{\mathrm{ZF}} x)]_{\mathrm{ZF}}\emptyset_{\mathrm{ZF}}\\
&\mathrm{ZF} \models [\lambda F\, \neg\exists y(Fy \,\&\, y \in_{\mathrm{ZF}} \emptyset_{\mathrm{ZF}})]_{\mathrm{ZF}} S_{\mathrm{ZF}}\\
&\mathrm{ZF} \models [\lambda R\, \neg\exists y(S_{\mathrm{ZF}}y \,\&\, y\, R\, \emptyset_{\mathrm{ZF}})]_{\mathrm{ZF}} \in_{\mathrm{ZF}}
\end{aligned}$$

But our question is, what reading should we assign to the unadorned ZF claim that "No set is an element of empty set"? Our answer is that we should assign the *conjunction* of the following three theorems:

$$\begin{aligned}
&\emptyset_{\mathrm{ZF}}[\lambda x\, \neg\exists y(S_{\mathrm{ZF}}y \,\&\, y \in_{\mathrm{ZF}} x)]_{\mathrm{ZF}}\\
&S_{\mathrm{ZF}}[\lambda F\, \neg\exists y(Fy \,\&\, y \in_{\mathrm{ZF}} \emptyset_{\mathrm{ZF}})]_{\mathrm{ZF}}\\
&\in_{\mathrm{ZF}}[\lambda R\, \neg\exists y(S_{\mathrm{ZF}}y \,\&\, yR\emptyset_{\mathrm{ZF}})]_{\mathrm{ZF}}
\end{aligned}$$

These were all proved to be theorems at the end of Section 4.1. We suggest that the *conjunction* of all three theorems captures the facts embodied by the unadorned mathematical claim "No set is an element of the empty set", for this latter claim is not only a fact about the empty set of ZF, but also a fact about the ZF-property of being a set and about the membership relation of ZF.

Indeed, we can, in general, *abbreviate* the conjunction of the above three theorems as a simple, intuitive formula:

$$\emptyset_{\mathrm{ZF}} S_{\mathrm{ZF}} \in_{\mathrm{ZF}} [\lambda xFR\, \neg\exists y(Fy \,\&\, yRx)]_{\mathrm{ZF}}$$

We can intuitively think of this as a 3-place encoding formula of the form $xyzH$, where the type of H is that of a three-place relation among entities having the types of x, y, z in that order. Of course, this isn't a well-defined formula of our logical framework,[17] but it doesn't need to be, for it simply serves as an abbreviation of a conjunction of well-formed formulas. In the particular case at hand, three places are needed because there are three primitive theoretical expressions in our target sentence ($\emptyset$, S, $\in$). But it is straightforward to define a function that takes as input an unadorned theorem of ZF and yields as output a n-place encoding formula of the above form, where the output encoding formula is an abbreviation of n well-formed encoding formulas. We omit the details here, though interested readers are directed to Zalta 2000 (250–251).

This, then, is a procedure for assigning a true (T-relative) reading to every theorem of an arbitrary mathematical theory T. It completes the reduction of mathematics to the above axiomatic system. In the next section we will show that each of the axioms of our system counts as logical or analytic.

We note here that the procedure and analysis described above offers a logical reconstruction of mathematical objects and relations as they are given antecedently by some concrete mathematical theory. While this provides, for example, a complete reconstruction of ZF-sets, we are not claiming that the reconstruction is necessarily a complete theory of sets since ZF might not be the complete theory of sets. Rather, we are offering the above as an analysis of any claim a mathematician might make in any context in which the mathematician is adopting all and only the assumptions of ZF. And this generalizes to other mathematical theories: in every context in which a mathematician assumes the principles of a mathematical theory T, we can use the above methods to analyze their claims.

5 Why This is Logicism

Logicism, historically, is the claim that every true mathematical proposition is derivable from the laws of logic extended with analytic truths.[18] Since we are focusing only on theoretical mathematics, logicism can be

[17]It might also be of interest to know that recent investigations of object theory start with primitive n-ary encoding formulas, and not just with unary encoding formulas. For purposes of this paper, we essentially defining n-ary encoding in terms of unary encoding.

[18]Some philosophers, e.g., Roeper 2015 and Klev 2017, take logicism to be the narrower claim that *arithmetic* is reducible to logic, but we take logicism to be more broadly conceived.

restated as the following clearer and simpler thesis: all mathematical theorems are derivable from the laws of logic extended with analytic truths, where by 'mathematical theorems' we mean any statement that is derivable from any axiomatic mathematical theory. If one prefers, one might paradigmatically identify the set of mathematical theorems simply with the theorems of ZFC, that is, first- or second-order Zermelo-Fraenkel set theory with the axiom of choice. But this is by no means required; mathematics may not have set-theoretic foundations, or it may not have mathematical foundations at all.

Moreover, in the Frege-Russell tradition, logicism consisted of an additional claim, to the effect that mathematical concepts are (analyzable in terms of) logical concepts.[19] Thus, we may understand logicism as the conjunction of the following two theses (Carnap 1931):

LC Logicism about Mathematical Concepts: Every mathematical concept denoted by a mathematical term is (identical to) a logical concept.

LP Logicism about Mathematical Theorems: For every mathematical theorem, if each term denoting a mathematical concept in any such theorem is replaced by a term denoting the logical concept identical to the mathematical concept, then the resulting theorem is a logical truth.

Clearly, as we have formulated these two principles, LP presupposes LC.

Logicism has traditionally been formulated primarily as a question of truth, and not logical consequence, since the emphasis has been on reducing mathematical claims to logical truths and not on showing that mathematical inferences are purely logical inferences. But we want our conception of logicism to extend to the idea that mathematical practice involves a body of inferences, so that logicism also encompasses the idea that mathematical truths can be derived as logical consequences of a logic (cf. Rayo 2005, 204). However, in what follows, we focus primarily on LC and LP and, along the way, note how our understanding yields logicism with respect to the notion of logical consequence defined below.

Our plan, then, is as follows:

[19]Here we shall be talking about *well-defined* mathematical concepts. We take the *well-defined* mathematical concepts of a theory T to be those represented by a term (i.e., an individual term or a predicate) of T that is either primitive or uniquely definable in T.

- First, we argue that the axioms presented in Section 3 are all either logical truths or analytic. In the case where we take the axioms to be logically true, we shall argue for their logicality by putting forward what we take to be a correct conception of logical truth, and then showing that under that conception, these axioms are logical truths. It is a consequence of our argument that the abstract objects picked out by our canonical descriptions are logical objects.

- Second, we argue that the additional axioms, put forward in Section 4 for the analysis of mathematical language, are all analytic.

- From the conclusions of these two arguments, we may then immediately conclude that the theorems of mathematics—represented in object theory as explained in Section 4.2—are logical or analytic truths, since theorems derived from logical and analytic truths are either logical or analytic. Moreover, we shall see (in Section 5.2) that a minor variant of LC follows as well, namely, that every mathematical concept denoted by a mathematical term is (identical to) a logical object.

5.1 Our Framework Axioms Are Logical or Analytic

In this section, we shall not argue, but rather assume, that the principles of classical logic, the substitution of identicals, the axiom governing descriptions, and the principle of λ-Conversion (i.e., the axioms discussed in Section 3 prior to Section 3.3) are logically true (as this notion is defined below) or analytic, where analyticity is defined in the usual way as truth in virtue of meaning (in this case, of the logical symbols). The principles of classical logic and the substitution of identicals have traditionally been regarded as logical. We add the law governing descriptions and λ-Conversion (AXIOMS 1 and 2) to this list of logical truths on the grounds that they are true in virtue of the meaning of the expressions *the* (represented by the $\imath$) and *being such that* (represented by the λ).[20]

[20]We recognize that λ-Conversion has existential import; one can derive existence claims from it, namely, the comprehension principle for relations. That is, one can derive $\exists F\forall x(Fx \equiv \varphi)$ from $[\lambda x\, \varphi]x \equiv \varphi$. But having existential import hasn't disqualified the laws of classical logic from being considered logical. The existential import of λ-Conversion is analogous to that of the quantifier laws; just as the latter imply that the domain of individuals must be non-empty, the principle of λ-Conversion implies that each higher-order domain is not just non-empty but has at least two properties, one that everything exemplifies and one that is empty. Both the quantifier laws and λ-Conversion make minor existence demands.

Moreover, AXIOMS 4–7 (in Sections 3.3.1 and 3.3.2) stipulate what is meant by the property of *being abstract* and, as such, are nothing more than meaning postulates. Once we take *abstract* objects to be those entities that are individuated by the properties they encode, these axioms articulate the conception of such objects in formal detail and, hence, are analytically true. Thus, we see the burden of the present paper as showing that the comprehension principle for abstract objects (AXIOM 3) is a logical truth, despite the fact that it baldly asserts the existence of abstract objects.

We begin with the observation that the classical understanding of the model-theoretic interpretation of the predicate calculus has overlooked one key feature of such interpretations. In particular, model-theoretic interpretations *should* include, in the domain of interpretation of the variables, everything that is required for the very possibility of predication, logically complex thought (including abstract mathematical thought), and logical consequences of those thoughts. That's the point of (a) thinking that the predicate calculus is a fundamental system for expressing our thoughts and validating inferences, and (b) thinking that an interpretation of that system will give us an insight into what's required for the possibility of having those thoughts and making those inferences.

To approach our thesis, let's reconsider why λ-Conversion is a logical truth. Consider one of its instances, which is logically complex not only because it involves the λ-expression but also because it involves the negation symbol:

$$[\lambda y\, \neg Gy]x \equiv \neg Gx$$

This holds for any property G: something exemplifies the negation of G iff it fails to exemplify G. There exists a logical pattern that underlies this fact, one that everything that fails to exemplify G has in common! After all, entities in the world do divide up into those that exemplify G and those that do not, and without the existence of the negation of G. How could two individuals a and b which both fail to be G not share the pattern of what is most-easily described as "being a non-G exemplifier"? If we treat this property of being a non-G exemplifier as reifying or representing this exemplification pattern, then it is required in any domain that contains the entities needed for truth of multiple predications of the form "x is a non-G exemplifier" ($[\lambda y\, \neg Gy]x$). And λ-Conversion also provides the *logical* justification as to why it is correct to infer one side of the biconditional from the other.

This same argument now applies to other instances of λ-Conversion, e.g., those involving other complex formulas such as conjunctions, disjunctions, conditionals, etc. The instances of λ-Conversion are true in any domain that contains the entities needed for the truth of predications involving complex predicates such as "x is G-and-H" ($[\lambda y\, Gy\, \&\, Hy]x$), etc., and provide the logical justification for inferences to and from such predicates, such as the inference from "x is G-and-H" to "x is G" (justifies $[\lambda y\, Gy\, \&\, Hy]x \vdash Gx$). The point also applies to relations and relational λ-expressions. Complex reasoning about the converse of relation R ($[\lambda xy\, Ryx]$)[21] and relations like unrequited love ($[\lambda xy\, Lxy\, \&\, \neg Lyx]$) assumes that the domain contains such relations. And, in general, the Comprehension principle for relations is logically true precisely because it postulates the entities that support such complex relational reasoning.

This leads us to a somewhat different philosophical conception of logical truth and logical consequence. We suggest: a sentence is *logically true* if and only if it is true in every interpretation that includes all the entities required for the possibility of thought, where this includes complex thought. Intuitively, logic is committed to the existence of whatever entities are required to support our inferences when reasoning about any (abstract) domain. We also say that a sentence φ is a *logical consequence* of a set Γ if and only if φ is true in every interpretation that (i) includes all the entities required for the truth of (logically complex) thought and (ii) makes every member of Γ true.

The Tarskian and Fregean conceptions of logical truth dovetail in our conception: from Tarski we take the idea that logical truth is truth in all models (Tarski 1936); from Frege we take the conception of logic as providing constitutive norms of thought as such (MacFarlane 2002), including constitutive norms for logically complex thought.[22] By combining the two, we end up with a definition of logical truth as truth in all models that include everything that is required for (logically complex) thought.[23]

[21] For example, arguing that a non-symmetric relation is distinct from its converse.

[22] For other views on Frege's conception of logic, see Goldfarb 2001, Linnebo 2003, and Blanchette 2013.

[23] This is consistent with the idea of logical truths as those that are constitutive of thought in general, and which are thus constitutively *a priori* in the sense discussed by Friedman 1994. (A sentence φ is constitutively *a priori* for a theory T just in case it is presupposed by T.) Indeed, one might perhaps think of the argument to be given in the present section as a kind of transcendental argument.

Under this conception, not only is λ-Conversion a logical truth, but the inference from $\neg Gx$ to $[\lambda y\, \neg Gy]x$, for example, is one of logical consequence.

Let's then see how this conception can be used to understand why the lowest level instances of AXIOM 3 (i.e., the instances asserting the existence of abstract individuals) are logically true. How could they be true in every model that includes everything that is required for (logically complex) thought? The early Greek mathematicians adopted a method that has persisted until this day: they attempted to characterize objects whose *only* properties are the properties given by the defining principles. Consider a simple example, such as discussions of The Equilateral Triangle. An early mathematician might have thoughts about an object that is triangular and has sides of equal length. In thinking about this object in the abstract, the mathematician might *logically infer* that The Equilateral Triangle is not scalene.[24] Here we have a simple logical conclusion in the form of a predication about The Equilateral Triangle and the domain must have such an object for the thought to be true and the inference to be valid.

Thus, the mathematician has defined, objectified, and drawn inferences about a pattern of properties of individuals. The assertion that this pattern exists as an individual is required for complex thoughts and inferences about The Equilateral Triangle. It is important to emphasize here that the existence of this pattern doesn't commit us to saying that there is an object that exemplifies the properties defining the pattern. In fact, we have two options that avoid this commitment but which offer an object of thought: either treat the pattern as a property of properties in 3rd-order logic, or treat it as an *abstract* individual that encodes the properties in question. Clearly, the appearance of the simple predication suggests that the latter is the way the mathematician has conceived of The Equilateral Triangle, namely, as an abstract individual. This justifies the logicality of comprehension over abstract individuals. And AXIOM 3 gives us the means to have logically complex thoughts about any other combination of properties that might be of interest, so that we may consider them as individuals in the abstract and reason about them. If one

[24] This example is representative of modern mathematicians as well. Consider Dedekind, who defined his simply infinite systems as consisting of objects whose *only* properties were those given by the axioms in his 1888 (§71).

is willing to accept λ-Conversion as logical, which objectifies patterns of exemplification, then one should be be willing to accept AXIOM 3 as logical, since it objectifies patterns of encoding in ways that are required for complex thoughts and what they logically entail. In other words, if one recognizes that second-order comprehension is logical because it merely expresses the existence of entities presupposed for higher-order thinking and reasoning, then one should also recognize that comprehension over abstract individuals is logical because it (analogously) merely expresses the existence of entities presupposed for the possibility of such activities.

Similar conclusions now apply to higher-level λ-Conversion and higher-level abstracta. For take the example:

$$[\lambda R\, \neg\forall x Rxx]S \equiv \neg\forall x Sxx$$

This asserts: relation S exemplifies being a non-reflexive relation iff S fails to be reflexive. There is a pattern of which every relation that fails to be reflexive is a part! Clearly, relations in the world do divide up into those that are reflexive and those that are not, and without the existence of the negation of the property of reflexivity $[\lambda R\, \neg\forall x Rxx]$, how could two relations S and S' which both fail to be reflexive not share the pattern of what is most-easily described as "being a non-reflexive relation"? If we treat this property of being a non-reflexive relation as reifying or representing this pattern, then it is required in any domain that contains the 2nd-level properties of relations needed for truth of multiple predications of the form "S is non-reflexive" ($[\lambda R \neg\forall x Rxx]S$). And so on to other instances of higher-order λ-Conversion. Thus, higher-order λ-Conversion is logical because it merely expresses the existence of higher-order properties like *being non-reflexive*, which are presupposed for a higher-order thought such as "S is non-reflexive" and for reasoning to the conclusion that a non-reflexive relation is not an equivalence relation.

And this is likewise the case for higher-level comprehension over abstract entities. Thus, let's consider why a particular instance of the comprehension principle for abstract relations, as applied to mathematical relations, is a logical truth. Consider the less-than relation ($<_D$) as given by the theory of dense linear orderings without endpoints. This relation is given by the following theory T_D:

$$\forall x, y, z(x < y \mathbin{\&} y < z \rightarrow x < z) \qquad \text{(Transitivity)}$$
$$\forall x(x \not< x) \qquad \text{(Irreflexivity)}$$
$$\forall x, y(x \neq y \rightarrow (x < y \vee y < x)) \qquad \text{(Connectedness)}$$

$$\forall x, y \exists z(x < z < y) \qquad \text{(Dense)}$$
$$\forall x \exists y \exists z(z < x < y) \qquad \text{(No Endpoints)}$$

The theorems derivable from these axioms constitute the theory T_D. What is more, we think of the $<_D$ relation itself as being constituted by that theory as well. The world itself doesn't contribute any facts about $<_D$ and there is no guarantee that a relation exists that exemplifies the properties of the $<_D$ relation — all there is to $<_D$ are the properties it has given this theory. In other words, the truths that ground all the facts about $<_D$ are facts of the form "In the theory T_D, $\mathcal{R}{<_D}$", where $\mathcal{R}$ ranges over properties of relations. That is, the theorems in the scope of the operator "In the theory T_D, ..." ascribe to relation $<_D$ various properties of relations, such as the properties of being transitive, irreflexive, connected, dense, and having no endpoints, and those that follow from them. There exists a pattern of predications, embedded within the theorems of T_D governing $<_D$, that we may articulate as a pattern of properties of relations, namely, the pattern: $T_D \models \mathcal{R}{<_D}$. $<_D$ just *is* that pattern of properties of relations, but instead of representing this pattern as a property of type $\langle\langle\langle i, i\rangle\rangle\rangle$ (property of properties of relations), encoding predication turns that pattern into an abstract relation of type $\langle i, i\rangle$ that encodes the properties of relations $\mathcal{R}$ that satisfy the pattern $T_D \models \mathcal{R}{<_D}$. (This is expressed in terms of our Reduction Axiom Schemata, discussed in Section 4.1.)

Indeed, that relation must exist for us to have a mathematical thought, and draw inferences, about the relation $<_D$. Hence, the notions of logical truth and consequence defined above have the following application: a sentence φ containing the term $<_D$ is *logically true* if and only if φ is true in all interpretations that include all the (abstract) objects required for the possibility of having the thought that φ. So $<_D$ must be contained in the domain of interpretation. Thus, the following existence claim is logically true:

$$\exists \textsc{r}(A!\textsc{r} \mathbin{\&} \forall\mathcal{R}(\textsc{r}\mathcal{R} \equiv T_D \models \mathcal{R}{<_D}))$$

This asserts: there is an abstract relation that encodes all and only the properties of relations exemplified by $<_D$ in T_D. And, generally, for any relation S of mathematical theory T, to have a thought about S, the following must be true:

$$\exists \textsc{r}(A!\textsc{r} \mathbin{\&} \forall\mathcal{R}(\textsc{r}\mathcal{R} \equiv T \models \mathcal{R}\textsc{s}))$$

Notice the theory T_D is a simple case in which the only distinguished term of the mathematical theory is a relation term. More complex mathematical theories involve both distinguished relation terms and distinguished individual terms. Thus, Peano Arithmetic has as primitives: the property *being a number*, the relation *successor*, and the individual *Zero*.

In such cases, the existence of the abstract property *being a number*$_{\text{PA}}$, of the abstract relation property *successor*$_{\text{PA}}$, and the abstract individual *Zero*$_{\text{PA}}$ are asserted by the relevant instances of comprehension AXIOM 3. The argument we developed above for justifying the logicality of AXIOM 3 for the type of individuals and the type of relations of individuals now generalizes to a justification of AXIOM 3 for the type of properties of individuals. That is, comprehension over higher-order abstract properties is also logically true, since it is required for the analysis of complex thoughts and inferences about mathematical properties. Arbitrary examples of higher-order instances of AXIOM 3, which assert the existence of abstract properties, simply assert the existence of something that is already in the relevant domain, namely, the patterns of properties of properties of individuals that already exist given any body of ordinary predications.

Thus, when our analysis is applied to Peano Arithmetic:

- There are at least *three* kinds of exemplification patterns that exist in the sentences prefaced by the operator "In Peano Arithmetic, ...", namely, patterns of properties of the property of *being a number*, patterns of properties of the *successor* relation, and patterns of properties of the individual *Zero*. (There are, additionally, patterns of properties of both the relations and properties definable in PA, but we'll discuss those below.)

- These particular patterns induce *three* corresponding kinds of encoding patterns that exist in the data of the form "In Peano Arithmetic, ...", namely, patterns of properties of the property of *being a number* in PA, patterns of properties of the *successor* relation in PA, and patterns of properties of the individual *Zero* in PA.

So, it follows from AXIOM 3 that the existence of the entities needed for the the analysis given by the Reduction Axiom Schemata is a logical truth.[25]

[25]This approach to logicism advances the ideas in Hodes 1984 (143) in several ways: his idea that the theory of natural numbers is an "encoding of a fragment of third-order logic" has been worked out in a systematic way, with the notion of *encoding* made rigorous. Moreover, we've applied the same technique to mathematical relations.

In this section, we have argued for the logicality of axioms that guarantee the existence of two general kinds of logical entities:

- Those which exist as exemplification patterns among individuals, properties and relations and which, by comprehension, are logical objects within the domain of (higher-order) properties, i.e., those whose existence is asserted by AXIOM 2.

- Those that exist as predication patterns (either exemplification or encoding patterns) among properties and relations, that, by comprehension for abstract individuals, comprehension for abstract properties, and comprehension for abstract relations, are logical objects within the respective domains, i.e., those whose existence is asserted by AXIOM 3.

Both kinds of entities are *logical* in so far as they are patterns of predications. The entities asserted to exist by AXIOMs 2 and 3 are *abstracted* from *pure logical patterns* formulable solely in terms of predications generally in our language. Given that the conditions under which they are asserted to exist correlate with pure logical patterns that exist in our language, what else could they be but logical objects?

Note that if our axioms are logical or analytic, then any *theorem* we can derive is logical or analytic.[26] Now if we can show that the claims of mathematics prefixed by the theory operator, which are imported when we apply object theory, are analytic, then it will follow that the *unprefixed* encoding claims of mathematics derived from the prefixed claims and object theory (at the end of Section 4.1)become logical or analytic. So we now turn to a defense of the idea that when we extend object theory in the application to mathematics, the new axioms are analytic.

[26]This point assumes that rules of inference preserve analyticity and logicality. This is clear in the case of analyticity. But we think it holds even of our new notion of logicality. We claim that if axioms φ and $\varphi \rightarrow \psi$ are logical in virtue of being required for the possibility of abstract thought and reasoning, then ψ is logical for the same reason.

5.2 The Additional Axioms for Mathematics are Analytic

Our goal in this section is to show that the axioms added to object theory in Section 4, namely, those introduced by the Importation Principle and the Reduction Axiom, are analytic (we don't need to justify the Closure Principle since it is just a special case of the Importation Principle). These are principles that underlie our analysis of mathematics.

We take it to be uncontroversial to claim that axioms introduced by the Importation Principle are analytic: we can put aside the controversial question of whether "$\emptyset \in \{\emptyset\}$" is analytic, and yet still claim that "In ZF, $\emptyset_{\mathrm{ZF}} \in_{\mathrm{ZF}} \{\emptyset\}_{\mathrm{ZF}}$" is. The latter is true in virtue of the meaning of the terms 'ZF', '$\emptyset_{\mathrm{ZF}}$', '$\in_{\mathrm{ZF}}$', and '$\{\emptyset\}_{\mathrm{ZF}}$'. Since 'ZF' denotes a theory, and a theory encodes its theorems, 'In ZF, $\emptyset_{\mathrm{ZF}} \in_{\mathrm{ZF}} \{\emptyset\}_{\mathrm{ZF}}$', when represented as $\mathrm{ZF} \models \emptyset_{\mathrm{ZF}} \in_{\mathrm{ZF}} \{\emptyset\}_{\mathrm{ZF}}$, is true in virtue of the meaning of all the terms used in the representation. (Henceforth we shall drop the subscripts for ease of reading.)

It remains to argue that axioms introduced by the Reduction Axiom are analytic. To do this, we argue that the meanings of the terms flanking the identity sign in instances of the Reduction Axiom are identical, i.e., that the meaning c_τ of a mathematical term τ is identical to the meaning of the canonical description used to identify the denotation of τ. As we shall see, this conclusion almost immediately implies a version of LC.

So, to argue that the instances of the Reduction Axiom are analytic, consider any mathematical concepts and the mathematical theories in which they occur. As examples, we again consider the concepts and the axioms of ZF set theory. Recall, first of all, that we formulated the following instances of the Reduction Axiom Schemata:

$$\begin{aligned}
&\emptyset_{\mathrm{ZF}} = \iota x(A!x \,\&\, \forall F(xF \equiv \mathrm{ZF} \models F\emptyset_{\mathrm{ZF}}))\\
&S_{\mathrm{ZF}} = \iota F(A!F \,\&\, \forall \mathcal{F}(F\mathcal{F} \equiv \mathrm{ZF} \models \mathcal{F}S_{\mathrm{ZF}}))\\
&\in_{\mathrm{ZF}} = \iota R(A!R \,\&\, \forall \mathcal{R}(R\mathcal{R} \equiv \mathrm{ZF} \models \mathcal{R}\in_{\mathrm{ZF}}))
\end{aligned}$$

The right-hand side of these identity statements involve canonical definite descriptions (we call them canonical T-based descriptions below). These descriptions are formulated with the new indexed terms introduced when the mathematical theories are imported into object theory.

By referencing these descriptions, we may give the following argument for the claim that the instances of our Reduction Axiom Schemata are

analytic (and once we establish that, we give an argument for a slight variant of our thesis LC of Logicism about Concepts). Let τ be any unambiguous mathematical term used in a mathematical theory T, where τ is either an individual term or a predicate of T:

(P1) The meaning of a mathematical term τ is a particular mathematical concept, which we shall denote as c_τ.

(P2) The mathematical concept c_τ is (identical to) the inferential role of τ in the theory T.

(P3) The inferential role of τ in the theory T is the object denoted by the canonical T-based description for τ, which is a logical object.

(P4) The logical object denoted by the canonical T-based description is also the meaning of the canonical T-based description.

So by transitivity of identity, the meaning c_τ of a mathematical term τ is identical to the meaning of the canonical T-based description for τ. And by the uncontroversial principle that if the meanings of τ and τ' are identical, then $\tau = \tau'$ is analytic, it follows that:

(A) The Reduction Axioms are analytic.

Here, then, is our support for the premises of the above argument.

We take (P1) to capture the intuitive way we think about mathematical terms and mathematical concepts. It makes explicit the pretheoretical understanding of these notions and should therefore be non-controversial.

Concerning (P2). We think it is reasonable to suppose that a mathematical concept c_τ is constituted by the systematic, as opposed to random, use of the mathematical term τ. In turn, the systematic use of τ can be grounded in a system of axioms and inferences in which that term appears. Thus, c_τ can be identified with the inferential role of τ in T. For example, $c_\emptyset$, c_π, $c_\in$, $c_<$, etc., are identical to the inferential roles of $\emptyset$, π, $\in$, $<$, etc., in their respective theories.

Of course, (P2) shouldn't be unfamiliar. It has an illustrious history. It can be understood as one way of spelling out Wittgenstein's meaning-as-use doctrine (1953), which found expression in later philosophers such as Sellars (1980) and Brandom (1998). (P2) is also an example of a view that is often found in the philosophy of science, tracing back to Schlick's and Carnap's view of theoretical terms in science. Both are influenced by

Hilbert's view of geometry, in which the meanings of mathematical terms are determined completely by the theories in which they figure.[27] This view of mathematical terms is even easier to accept than the corresponding view of theoretical terms in science, given that unlike the latter, the mathematical terms aren't necessarily introduced with the idea of representing some empirical entity. Finally, (P2) is consistent with Frege's Context Principle, except that the meaning for a mathematical term is not given by any single reference-fixing sentence but rather by a whole theory.

Concerning (P3). We think this can be seen by considering examples, such as the following one: the inferential role of '$\emptyset$' in ZF is properly identified by the canonical description, $\imath x(A!x \;\&\; \forall F(xF \equiv \textsc{zf} \models F\emptyset_{\mathrm{ZF}}))$. This holds because the abstract object denoted encodes all and only the properties of the null set derivable in ZF.[28] For instance, it is derivable in ZF that $\emptyset \in \{\emptyset\}$. This theorem gets imported into object theory as (suppressing subscripts) $\textsc{zf} \models \emptyset \in \{\emptyset\}$. So by the Closure Principle, it follows that $\textsc{zf} \models [\lambda x\; x \in \{\emptyset\}]\emptyset$. As such, the property $[\lambda x\; x \in \{\emptyset\}]$ is one of the properties that satisfies the formula $\textsc{zf} \models F\emptyset_{\textsc{zf}}$ in the matrix of the canonical description $\imath x(A!x \;\&\; \forall F(xF \equiv \textsc{zf} \models F\emptyset_{\textsc{zf}}))$. In object theory, the inferential role of the symbol $\emptyset$ is constituted by the object that encodes the totality of such properties. As such, the inferential role of this term is properly identified by our canonical T-based description.

Concerning (P4). Our argument for this premise begins with the inspection of the semantics of our formalism, which reveals that the terms of our formalism are assigned only one semantic value. We claim that this semantic value serves both as the denotation and meaning of the terms of our formalism. Our semantics assigns meanings by assigning denotations.

[27]See, for example, Friedman 1999, where we find (26):

> In *General Theory of Knowledge*, his [Schlick's] starting point is Hilbert's *Foundations of Geometry* and the notion of axiomatic or *implicit* definition [...] According to the conception that Schlick derives from Hilbert, the primitive terms of geometry require no intuitive meaning or content. All we need to know about these primitives for the purposes of pure geometry are their mutual logical relationships set up explicitly in the axioms. Points, lines, and planes are any system of objects whatsoever that satisfy these axioms.

[28]Again, our analysis is one way of developing a *proof-theoretic semantics*, since we are generating term meanings by abstracting over the proof-theoretic roles of the relevant terms. See Prawitz 2006, Schroeder-Heister 2006, Francez & Dyckhoff 2010.

Indeed, we take it that *for our formalism*, the distinction between the denotation and meaning of its terms just collapses.[29]

Some philosophers have thought that the intension/extension distinction that applies to natural languages must also apply to formal languages and that to properly represent the intensions and extensions of the terms of natural language, one must build a formal language with terms having both intensions and extensions. (One can construe systems by Carnap 1947, Church 1951, and Montague 1973 in this way.) But we reject both suggestions. It is not obvious that the intension/extension distinction must be applied to formal languages. The method of denotation and satisfaction, derived from Tarski (1933) and noted by Myhill (1963), suffices to give a semantics for the terms of a formal language without introducing intensions and extensions. Moreover, one doesn't have to build a formal language with terms having both intensions and extensions in order to model the intensions and extensions of the terms of natural language. One simply needs to have (a) terms in the formal language that can represent the extensions of the terms of natural language as well as (b) terms in the formal language that can represent the intensions of the terms of natural language. That is what our system does.[30]

Given P1 – P4, therefore, we've established (A), i.e., that the instances of our Reduction Axiom are analytic. This completes our argument that the additional axioms added to object theory for the analysis of mathematics are analytic, and so we have established logicism in the form

[29]As a consequence of this fact, it doesn't matter whether we say that a term "denotes" or "expresses" its semantic value.

[30]In the Appendix, we build the smallest *extensional* model of object theory. For purposes of showing consistency, this suffices. But one can build models in which the denotation of an n-place predicate is not just a set (or a truth-value, in the case of a 0-place predicate). Indeed, we take it that in the intended models, n-place predicates denote n-place relations, where the latter are then systematized by the principles for n-place relations offered by object theory (i.e., the principles laid down in Sections 3.2.2 and the definitions for the identity of relations given in Section 2). The result will still be an 'extensional' model but with primitive 'intensional' entities or urelements populating the relational domains. Moreover, we need not represent abstract individuals and abstract relations in these more fine-grained models as sets of properties intensionally-conceived. Instead, they too, may be regarded as intensionally-conceived urelements, indeed, hyperintensionally-conceived urelements, given that they correlate with sets of intensionally-conceived properties. So the denotations of our canonical T-based descriptions may themselves be conceived as highly intensional entities. We see no need to add a second semantic value (i.e., an intension) for the terms of our formal language.

of LP. Since the theorems of mathematics are correctly representable as theorems of extended object theory, and the axioms of extended object theory are all either logical truths or analytic truths, then the theorems of mathematics are themselves either logical or analytic. This of course assumes that the rules of inference in classical logic preserve logicality and analyticity. We shall not argue for this claim.

Furthermore, from premises P1 – P3, a form of LC follows, again by transitivity of identity:

LC′ The meaning of a mathematical term τ, i.e., the mathematical concept c_τ, is identical to a logical object.

We take it that LC′ serves as a reconstruction of the logicist claim LC. Of course, if one is further convinced that our logical objects are indeed logical concepts, then we may derive LC itself. But we shall not argue for this, since we take it that LC′ is sufficient for our defense of logicism.

6 Objections and Observations

6.1 Objections

One objection that might be raised is whether we have offered an analysis that does 'too much', in that it would give us a means of reducing theoretical terms in natural science to logic! The objection argues that our very same procedure, as outlined above, would give us denotations for theoretical terms like 'electron', namely, the abstract property that encodes exactly the properties of properties attributed to this property by our best available physics. But, here, we argue, there is a disanalogy, which prevents one from properly applying the above analysis to theoretical terms of natural science. The disanalogy is that in natural science, the theoretical properties like being an electron are natural properties, whereas the theoretical properties of mathematics are not. Thus, in the case of natural science, one might distinguish the natural property from our various concepts of that property, as these concepts change from scientific theory to scientific theory. The property of being an electron, for example, is something there in the world, though our theories of the electron reflect our evolving concept of this property. The concept, but not the property, is tied to the inferential role.

Given this distinction, we would argue that our analysis above could not be applied to analyze the property of being an electron (though it might be applied to the concept *electron* as this might be embodied by some scientific theory). Thus, P1 fails in the case of the natural properties of physics: "the meaning of theoretical term τ in a physical theory" is a natural property, not a physical concept. Hence P1 is false. Whereas the physical concept might well be identical to an inferential role, as (the corresponding version of) P2 would have it, the physical property is not an inferential role at all. By contrast, in the case of mathematical properties, there is no distinction to be drawn between our concepts of a mathematical property and the property itself. Either the mathematical properties and our concepts of them collapse, since the former are not given by anything over and above the concepts, or there are no mathematical properties beyond our mathematical concepts. In the former case, we use the above analysis to identify both the property and the concept, collapsing the two; whereas in the latter case, we use the above analysis solely for understanding our mathematical concepts (in which case there is nothing else to understand).

Another objection might run as follows. Our analysis assumes that mathematical objects are identified in terms of actual theories, i.e., theories that someone has actually developed or asserted. Doesn't this imply that the abstract realm of mathematical objects depends on the contingent actions of humans? To this, we may reply that by showing how all axiomatically developed mathematics consists of logical/analytic truths, we have shown a striking fact that achieves the goals of logicism. But a deeper response is also available, since the objection suggests that the theorems of mathematics are contingent claims.

In fact, they are not. To see why, note that we've analyzed the theorems of mathematics as encoding truths about the individuals and relations of mathematics. Though we didn't develop the modal version of object theory here, one modal principle included in the theory is the claim that $\Diamond xF \rightarrow \Box xF$, i.e., that if possibly an abstract object encodes a property, then it does so necessarily.[31] So, though it may be a contingent fact as to what mathematicians have asserted by way of mathematical axioms, the theory-prefixed claims of the form "In theory T, p" are not contingent; they are analytic truths and given claims of that form, neither

[31] See the modal applications of object theory beginning with Zalta 1983, Chapter III.

the theorems of T (as we've analyzed them) are contingent claims nor are the objects of the theory contingent objects.

Moreover, note that our analysis extends to *any possible* mathematical theory. If we consider mathematical theories to be mathematical situations (i.e., abstract objects that encode only propositional properties constructed from mathematical propositions) that in fact have an author, then we can define a *possible mathematical theory* as any mathematical situation that possibly has an author.[32] Then we can say that our analysis applies not only to actual mathematical theories but also possible mathematical theories. Of course, we cannot import the axioms of those possible mathematical theories into object theory until a mathematician actually asserts them and turns them into actual mathematical theories, but the fact is that a possible mathematical theory T would have axioms and its axioms would be subject to the analysis developed in Section 4.

Thus, the realm of mathematical objects is not so closely tied to the contingent actions of human mathematicians. Though we haven't developed *modal*, 3rd-order object theory in this paper, it would be trivial to add a modal operator. If we adopt the axioms for S5 modal logic, then these possibility claims about possible authors are in fact necessary, and thus the realm of mathematics becomes defined on our view in terms of objects with no air of contingency about them.

Finally, a Platonist might object that when we extend a theory like ZF to ZFC, the mathematician is not talking about a different realm of sets, while our approach implies that they are. But in fact that is not the case: our approach is consistent with assigning "the" right denotations to set-theoretic terms and predicates and that perhaps these denotations are only incompletely described by both ZF and ZFC. Note that the Platonist claim presupposes both that sets exist independently of our theories of them and that when we move from ZF to ZFC, the new theory is simply characterizing the objects of ZF further. So, for the sake of argument, suppose that the sets do exist independently of our theories of them and that there is consequently a complete body of all set-theoretic truths. Introduce a proper name, say '$\mathfrak{G}$', for that body of truths and replace 'ZF' in our reduction axioms from above by '$\mathfrak{G}$'. Then everything should

[32]Of course, we are not attempting to define what is *mathematical* in this paper. We are presupposing that mathematicians can recognize what is mathematical and what is not. All we are attempting to do is analyze the language and theories that they claim to be mathematical.

go as before, and we should be able to reconstruct the mathematical terms of set theory as logical expressions. That is, we can plug $\mathfrak{G}$ into the machinery that we described above and the result, we claim, is a logicist reconstruction of *the* concept of set. Given this reconstruction, the theorems of ZF and ZFC will be true of the denotations that we have assigned to the terms and predicates of set theory $\mathfrak{G}$ since $\mathfrak{G}$ includes these theorems (assuming Choice is included in $\mathfrak{G}$). Of course, this is all completely hypothetical: we know that if *the* body of truths of set theory exists, it is not recursively axiomatizable, so we will never be "given" that body of truths in the form of a complete axiomatic system; nor does there seem to be any alternative manner in which we could be "given" that body of truths in a literal sense. But that does not affect the principal logicist point that we want to make.

6.2 Comparison with Other Approaches

At the present time, we know of no other successful version of logicism, i.e., no successful attempt to establish LC and LP. While Frege's version of logicism failed due to the inconsistency of his Basic Law V, Whitehead and Russell's account of logicism was based on principles, such as the Axiom of Reducibility and the Axiom of Infinity, whose status as logical truths were unclear at best. Similarly, efforts by Hodes (1991) and Tennant (1987) both require an appeal to non-logical, or at even, non-analytic, axioms of infinity,[33] and Tennant (1987) has to extend logic each time he analyzes a new mathematical object, by adding Introduction and Elimination rules to logic to govern the terms for the new mathematical objects.

In recent years, neologicist theories have been developed that rely on abstraction principles.[34] These neologicist theories add new abstraction principles for each new kind of mathematical object introduced and each of these new 'double abstraction-identity principles' (like Hume's Principle, which introduces two abstractions, $\#F$ and $\#G$ in the same principle) *combines* both a comprehension (or existence) claim and an identity claim

[33]So our worry concerning such axioms of infinity is not due to them postulating the existence of some kind of infinite object—which would be fine, as far as we are concerned, as long as the object in question is a logical object—rather what we worry is about is whether one can argue that these axioms of infinity are logical or, at the very least, analytic.

[34]See the work of Wright 1983, Hale 1987, Boolos 1985, Cook 2003, etc., and for an overview, see Linsky & Zalta 2006.

for the new kind of mathematical object. Clearly, any reduction of mathematics to logic will have to use some definitions or principles for identifying the mathematical objects as logical objects. On our view, however, only a single comprehension principle for objects is needed and, moreover, a single identity principle for objects is specifiable independently of comprehension.

Our approach differs from previous approaches in the following ways. (a) We appeal only to principles that are arguably logical or analytic and, in particular, we don't appeal to any non-logical axiom of infinity. Our unapplied and purely logical theory still has a finite model (described in the Appendix). When we apply the theory to mathematics, we sometimes import an infinite number of theorems into our own theory in the form of analytic truths. The infinity of mathematical entities that results from this extension is a presupposition of mathematical thought and thus counts as logical in our understanding of the term. (b) We don't have to continually re-prove our system is consistent since we don't add a new principle or rule of inference for each type of mathematical object; we use a uniform method for analyzing every kind of mathematical object. The model we have proposed in the Appendix—even though it is merely a minimal model for our background logical theory—grounds our conjecture that no special steps need be taken to guarantee consistency each time a new part of mathematics is analyzed in the manner outlined in Section 4. This difference looms especially large when one considers that with neologicism, the 'bad company' and 'embarassment of riches' objections force a neologicist to prove their system consistent each time a new abstraction principle is added to analyze some part of mathematics.[35] (c) Our approach is not subject to any of the traditional objections to the neologicist approach, such as the Julius Caesar problem,[36] the bad

[35] For the bad company objection, see Boolos 1990 (214), Field 1989 (158), and Dummett 1991 (188-189). For the embarassment of riches objection, see Heck 1992 and Weir 2003 (16). For recent work on the bad company objection, see the papers in Linnebo (ed.) 2009. See MacBride 2003 and Linnebo 2009 for a summary of the problem.

The neologicist could adopt some constraints, such as those proposed in Fine 2002, and suggest that any abstraction principle meeting those constraints would be 'safe' and not require a new consistency proof. But as far as we can tell, by placing those constraints on abstraction principles, one couldn't then derive all parts of mathematics from Fregean biconditionals. See Burgess 2003, 2005.

[36] This is the problem that Frege himself raised for his own view: when abstraction principles like Hume's Principle ($\#F = \#G \equiv F \approx G$) are added to second order logic

company objection, the embarassment of riches objection, etc. (d) Our analysis is prepared even for not-yet-formulated mathematical theories and new kinds of mathematical objects. Finally, (e) our approach gives an account of the denotations of both the individual terms as well as the predicates of mathematical theories.

If MacFarlane (2002) is correct, then the principles of object theory are *general* in the sense accepted by both Kant and Frege, namely, they are constitutive of, and provide a norm for, the possibility of having complex logical thoughts, including abstract mathematical thought. That is very different from the modern conception of logic, since our conception allows that some existence claims can be logical truths. Indeed, we suggest that the argument (in the previous section) for the logicality of both of our comprehension principles (one for relations of every higher-order type and one for abstract objects of every type) justifies the early logicist view that logic may endorse existence claims, namely, those that assert the existence of the logical objects that Frege, Russell, and Whitehead used to reduce mathematics. The only existence claims logic is committed to are those required for the possibility of having complex logical thoughts.

So logic does have ontological commmitments, but it commits one to nothing more than what is required for the possibility of formulating and interpreting complex predications. In particular, unapplied object theory has ontological commitments (see the model in the Appendix), but the unapplied theory is not committed to anything more that what is required for the possibility of reifying structural relations among relations, i.e., what is required to make sense of the abstract relations that emerge from patterns of exemplification predication that are available in first- and second-order logic.

As we've mentioned, our understanding of logic and logicality has consequences for certain controversies concerning existence claims. Logicians have faced the following issue: what should one say about the fact that standard first-order logic entails existence claims (such as $\exists x\, x = x$). This has traditionally been seen as an uncomfortable conclusion since it was thought, on the one hand, that logic should be free of existential commitment, whereas on the other hand, that first-order logic should be the

as the basis for identifying the numbers, identity is given only when two numbers are given in the form $\#F$ and $\#G$. The condition '$x = \#F$' is left undefined, and so the analysis yields no answer to questions like, "Is Julius Caesar identical to the number of Fs". In our system, '$x = \#F$' is always defined, since '$x = y$' is defined for every x, y.

background system for the assertion of any non-logical existence principles. However, we can comfortably accept the fact that logical axioms imply existence claims because it is now clear that a non-empty classical domain should come pre-stocked with a minimum group of abstract objects.

Moreover, some philosophers have argued that $\exists x \exists y(x \neq y)$ can't be a logical truth, since it asserts the existence of more than a non-empty domain. Our arguments in Sections 5.1 and 5.2 show that this formula is a logical truth.[37] We argued that some formulas are in fact logically true in the sense that they are true in all models that make it possible to have logically complex thoughts. Moreover, once we import mathematics into our system in the form of analytic truths, then we can derive the existence of new logical objects.

6.3 Epistemology Redux

We claim, finally, that the epistemological benefits of logicism now accrue. By showing that mathematical statements are analytic, it follows that by knowing the meanings of (the terms in) these statements, we are equipped with all the tools we need to determine whether they are true. We can know mathematical theorems by deriving them solely from logically true statements and analytic statements. We know the logically true statements on the grounds that they are part of the above foundations for logic (by formulating the above system, the logical truths in our axiomatic system are recursively axiomatizable – we know what the axioms are and we know how the rules of inference allow one to derive theorems from the axioms) and we know the analytic statements in virtue of their meaning alone. Thus, no special cognitive faculty for knowledge of mathematical truths is needed other than the faculty of understanding, which is a faculty we, like Benacerraf (1981), take to be explainable in naturalistic terms.

So we don't have to posit a causal information pathway, like the causal theory of reference, to explain how we come to understand the terms of mathematical statements. Our comprehension principles already constitute the paths by which we apprehend abstract objects: from the body of mathematical theorems of T, the comprehension principles just are the means by which we cognitively grasp the objects denoted by the terms of T. This goes back to the point in Linsky and Zalta 1995, in which it is argued that in the case of mathematics, knowledge by acquaintance and knowledge by description collapse: all that one has to do to become cognitively acquainted with a mathematical object or relation is to *understand* the canonical description that identifies it. Thus, we can determine the truth of a mathematical statement simply by what it says and reflecting on the properties encoded by the canonical descriptions identifying the denotations of its terms.

Of course, this might be a very difficult thing to do. It can be very hard to know the (full) meaning of a mathematical statement, because the objects and relations denoted by the terms of the statement are defined relative to the entire body of theorems in the theory in which the statement is made. The claim that mathematical truths are logical or analytic truths does not entail that for each mathematical truth it would be easy to determine that it is a mathematical truth. But we trust that this result is already inherent in logicism.

Moreover, what we have *not* done in the present paper is to say anything about the *a priori* justification of the logic underlying object theory. Our work thus far only shows that by reducing mathematics to a logic of the kind described above, our knowledge of mathematics is *a priori*, albeit relative to the *a priori* justification of our knowledge of the logical system to which it has been reduced. We would add to Benacerraf's point that that our knowledge of mathematics can be accounted for by whatever accounts for our knowledge of language *and* logic.

To be clear, though, the epistemological situation is very different from that surrounding the foundational system of *Principia Mathematica*. In the early part of the 20th century, philosophers and logicians would have agreed that the logic in that work is justified if the axioms of reducibility and infinity were in fact logical. But Whitehead & Russell couldn't very well argue that the axioms of reducibility and infinity are *logical*, even if they had tried to use the grounds we provided above: it is hard to see how such axioms are required for the possibility of (abstract) thought. Our logical framework, by contrast, requires no such axioms and the axioms it does assert are required for such thought. But in the present case, even if one accepts that our axioms are logical, there is still the question of whether they are justified. We have not addressed that latter ques-

[37] Note that $\exists x \exists y(\neg A!x \,\&\, \neg A!y \,\&\, x \neq y)$ is not a necessary truth – this isn't required for logically complex thought. But $\exists x \exists y(A!x \,\&\, A!y \,\&\, x \neq y)$ is and, thus, so is $\exists x \exists y(x \neq y)$, since the latter is implied by the former.

tion. We could argue that a logic such as ours is justified because it is *presupposed* somehow, or because the logic, through a process of reflective equilibrium, offers a rational reconstruction of the data (i.e., logical consequences we accept pretheoretically) that is better than other logical systems. But we have to leave this argument for another paper.

Appendix: A Minimal Model of the Logical Framework

The logic in the foregoing is consistent, as can be demonstrated by the construction of the smallest model. This model happens to be an extensional one: ordinary properties and ordinary relations are not distinguished from their exemplification extensions, and ordinary propositions are not distinguished from the two truth values The True and The False. Of course, this extensional model is not the intended one.[38]. We emphasize, however, that our model doesn't require "full higher-order semantics"; we don't require, of any higher-order domain, that it be the full power set of the lower order domain.

A Bounded Language

Our analysis of mathematics does not require the full unbounded language defined in the paper in Section 2. So, as we develop a model of typed object theory, we restrict our attention to the fragment we need. We shall therefore define a bounded language $\mathcal{L}_{n,m}$, where n and m are bounds that set, respectively, the *width* and *height* of the types for the expressions of the language. We begin by defining the functions h and w for the height and width, respectively, of a given type.

The *width* of type t, written $w(t)$, is defined as:

- $w(i) = 1$
- $w(\langle\,\rangle) = 1$
- $w(\langle t_1, \ldots, t_k\rangle) = \sum_1^k w(t_k)$

[38] An intended model would distinguish properties and relations from their exemplification extensions, and would distinguish propositions from their truth values. Thus, an intended model would be intensional, and if modality were added, would be hyperintensional.

The *height* of type t, written $h(t)$, is defined as:

- $h(i) = 0$
- $h(\langle\,\rangle) = 1$
- $h(\langle t_1, \ldots, t_k\rangle) = 1 + \max\{h(t_1), \ldots, h(t_k)\}$

Then we define $\mathcal{L}_{n,m}$ as the language that includes any well-formed expression of $\mathcal{L}$ that can be formulated only with terms τ of type t such that $w(t) \leq n$ and $h(t) \leq m$.

Before we define a model for the bounded language $\mathcal{L}_{n,m}$, a few observations are in order. Intuitively, we want to choose bounds that will yield the smallest language and model needed for our analysis. The following two considerations play a role in setting the bounds on $\mathcal{L}_{n,m}$:

- In Section 4.1, we analyze the property of being a ZF-set as an abstract property having type $\langle i\rangle$. Abstract properties of this type encode properties of type $\langle\langle i\rangle\rangle$. And properties of this latter type can be both ordinary and abstract, though the mathematical properties will encode only abstract properties of type $\langle\langle i\rangle\rangle$. But these abstract properties must, in turn, encode properties of type $\langle\langle\langle i\rangle\rangle\rangle$. Note this requires the bound on the height of the types in the language to be at least 3 and the bound on the width to be at least 1.

- In Section 4.1, we analyze the membership relation of ZF as an abstract relation having type $\langle i, i\rangle$. Abstract relations of this type encode properties of type $\langle\langle i, i\rangle\rangle$. And properties of this latter type can be both ordinary and abstract, though the membership relation will encode only abstract properties of type $\langle\langle i, i\rangle\rangle$. But these abstract properties must, in turn, encode properties of type $\langle\langle\langle i, i\rangle\rangle\rangle$. Note this requires the bound on the height of the types in the language to be at least 3 and the bound on the width to be at least 2.

Given these facts, it should be clear that the minimal fragment our analysis of mathematics requires is the bounded language $\mathcal{L}_{2,3}$. In line with what we said above, $\mathcal{L}_{2,3}$ includes any well-formed expression of $\mathcal{L}$ that can be formulated only with terms τ of type t such that $w(t) \leq 2$ and $h(t) \leq 3$. So in specifying a general model for $\mathcal{L}_{n,m}$, we shall occasionally focus on the model for the language $\mathcal{L}_{2,3}$.

The Smallest, Extensional Model for $\mathcal{L}_{n,m}$

We construct our model in two basic stages: first we construct the *structural* domains of the model, and second, we specify the *domains of quantification* and a proxy function (that assigns to each element in a domain of quantification to an element of a structural domain). The construction of the *structural* domain occurs in two stages: (1) the kernel of each type and (2) the abstract objects of each type.

Structural Domains: Kernel

(1) We define the *kernel* $\mathbf{K}_t$ of objects of type t, by induction, as follows:

- Where $t = i$, the kernel $\mathbf{K}_i$ of individuals is the union of two subdomains: the ordinary individuals $\mathbf{O}_i$ and the special individuals $\mathbf{S}_i$. For the purposes of building a specific minimal model, we stipulate that $\mathbf{O}_i$ is empty and $\mathbf{S}_i$ contains a single special individual $\boldsymbol{s}_i$, which we henceforth label as $\boldsymbol{b}$.

- Where $t = \langle\, \rangle$, the kernel $\mathbf{K}_{\langle\,\rangle}$ of propositions is the union of two subdomains: the ordinary propositions, $\mathbf{O}_{\langle\,\rangle}$ and the special propositions $\mathbf{S}_{\langle\,\rangle}$. For the purposes of building a specific minimal model, we stipulate that $\mathbf{O}_{\langle\,\rangle}$ contains two propositions, labeled $\mathbf{T}$ and $\mathbf{F}$, and $\mathbf{S}_{\langle\,\rangle}$ contains a single special proposition $\boldsymbol{s}_{\langle\,\rangle}$, which we henceforth label as $\boldsymbol{a}$.

- Where $t = \langle t_1, \ldots, t_n \rangle$ ($n \geq 1$), for any types $t_1, \ldots, t_n$, the kernel $\mathbf{K}_{\langle t_1,\ldots,t_n\rangle}$ of relations among objects having types $t_1, \ldots, t_n$, respectively, is the union of two subdomains, $\mathbf{O}_{\langle t_1,\ldots,t_n\rangle}$ and $\mathbf{S}_{\langle t_1,\ldots,t_n\rangle}$, where $\mathbf{O}_{\langle t_1,\ldots,t_n\rangle} = \wp(\mathbf{K}_{t_1} \times \ldots \times \mathbf{K}_{t_n})$ and $\mathbf{S}_{\langle t_1,\ldots,t_n\rangle}$ contains at least one special object $\boldsymbol{s}_{\langle t_1,\ldots,t_n\rangle}$. For the purposes of building a specific minimal model, we label:

  $\mathbf{s}_{\langle i\rangle}$ as $\boldsymbol{c}$
  $\boldsymbol{s}_{\langle i,i\rangle}$ as $\boldsymbol{e}$
  $\boldsymbol{s}_{\langle\langle i\rangle\rangle}$ as $\boldsymbol{m}$
  $\boldsymbol{s}_{\langle\langle i,i\rangle\rangle}$ as $\boldsymbol{n}$.

Given these stipulations, we have the following consequences:

- $\mathbf{K}_{\langle i\rangle}$ (i.e., the kernel of objects of type $\langle i\rangle$) $= \wp(\mathbf{K}_i) \cup \{\boldsymbol{c}\}$, and so $\mathbf{K}_{\langle i\rangle} = \{\, \{\boldsymbol{b}\}, \{\,\}, \boldsymbol{c}\,\}$. (This is pictured in the graphic.)

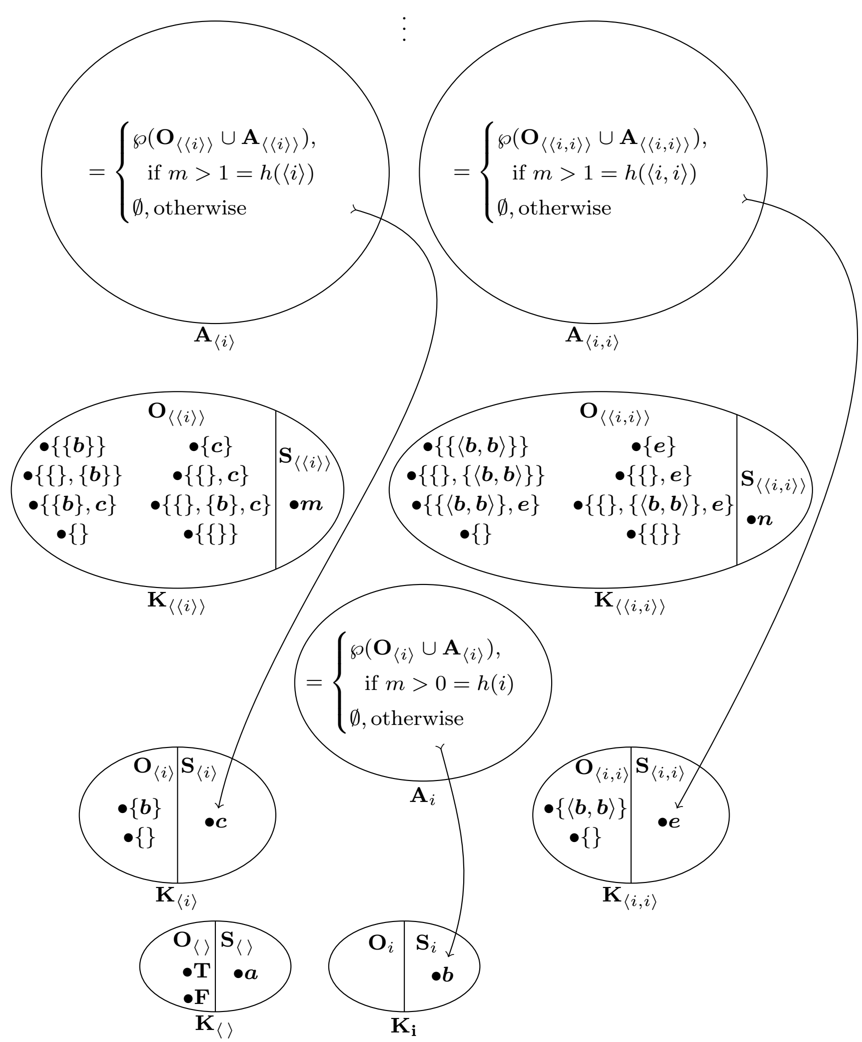

Figure 1: A fragment of the minimal model with unrestricted typed comprehension for abstracta. The domains, from the bottom up, are: the *kernel of individuals* $\mathbf{K}_i$ ($= \mathbf{O}_i \cup \mathbf{S}_i$); the *kernel of propositions* $\mathbf{K}_{\langle\,\rangle}$ ($= \mathbf{O}_{\langle\,\rangle} \cup \mathbf{S}_{\langle\,\rangle}$); the *kernel of properties of individuals* $\mathbf{K}_{\langle i\rangle}$ ($= \mathbf{O}_{\langle i\rangle} \cup \mathbf{S}_{\langle i\rangle}$); the *kernel of binary relations among individuals* $\mathbf{K}_{\langle i,i\rangle}$ ($= \mathbf{O}_{\langle i,i\rangle} \cup \mathbf{S}_{\langle i,i\rangle}$); the abstract individuals $\mathbf{A}_i$ (= the power set of $\mathbf{O}_{\langle i\rangle} \cup \mathbf{A}_{\langle i\rangle}$); the *kernel of properties of properties of individuals* $\mathbf{K}_{\langle\langle i\rangle\rangle}$; *the kernel of properties of relations among individuals* $\mathbf{K}_{\langle\langle i,i\rangle\rangle}$; *abstract properties of individuals* $\mathbf{A}_{\langle i\rangle}$ (= the power set of $\mathbf{O}_{\langle\langle i\rangle\rangle} \cup \mathbf{A}_{\langle\langle i\rangle\rangle}$); and *abstract relations among individuals* $\mathbf{A}_{\langle i,i\rangle}$ (= the power set of $\mathbf{O}_{\langle\langle i,i\rangle\rangle} \cup \mathbf{A}_{\langle\langle i,i\rangle\rangle}$).

- $\mathbf{K}_{\langle i,i\rangle}$ (i.e., the kernel of objects of type $\langle i,i\rangle$) $= \wp(\mathbf{K}_i \times \mathbf{K}_i) \cup \{\boldsymbol{e}\}$, and so $\mathbf{K}_{\langle i,i\rangle} = \{\,\{\langle \boldsymbol{b},\boldsymbol{b}\rangle\}, \{\,\}, \boldsymbol{e}\,\}$. (This is pictured in the graphic.)

- $\mathbf{K}_{\langle i,\langle\,\rangle\rangle}$ (i.e., the kernel of relations between individuals and propositions) $= \wp(\mathbf{K}_i \times \mathbf{K}_{\langle\,\rangle}) \cup \{\boldsymbol{s}_{\langle i,\langle\,\rangle\rangle}\}$, and so $\mathbf{K}_{\langle i,\langle\,\rangle\rangle} = \{\,\{\langle \boldsymbol{b},\mathbf{T}\rangle\}, \{\langle \boldsymbol{b},\mathbf{F}\rangle\}, \{\langle \boldsymbol{b},\boldsymbol{a}\rangle\}, \{\langle \boldsymbol{b},\mathbf{T}\rangle, \langle \boldsymbol{b},\mathbf{F}\rangle\}, \{\langle \boldsymbol{b},\mathbf{T}\rangle, \langle \boldsymbol{b},\mathbf{a}\rangle\}, \{\langle \boldsymbol{b},\mathbf{F}\rangle, \langle \boldsymbol{b},\mathbf{a}\rangle\}, \{\langle \boldsymbol{b},\mathbf{T}\rangle, \langle \boldsymbol{b},\mathbf{F}\rangle, \langle \boldsymbol{b},\mathbf{a}\rangle\}, \{\,\}, \boldsymbol{s}_{\langle i,\langle\,\rangle\rangle}\,\}$. (This is not pictured in the graphic.)

Etc.

Structural Domains: Abstract Objects

(2) Given m as the maximum height, we recursively define the domain $\mathbf{A}_t$ as follows, where $\boldsymbol{a}^t_\emptyset$ names a distinguished null abstract object of type t, i.e., the abstract object of type t that encodes no properties:

$$\mathbf{A}_t = \begin{cases} \wp(\mathbf{O}_{\langle t\rangle} \cup \mathbf{A}_{\langle t\rangle}) & \text{if } h(t) < m \\ \emptyset & \text{otherwise.} \end{cases}$$

The identity conditions for elements of $\mathbf{A}_t$ depend on the higher-type elements of $\mathbf{O}_{\langle t\rangle}$. Given these stipulations, we have the following consequences for $\mathbf{A}_{\langle i\rangle}$ and $\mathbf{A}_i$ when $m=1$:

$$\mathbf{A}_{\langle i\rangle} = \emptyset$$

$$\begin{aligned}\mathbf{A}_i &= \wp(\mathbf{O}_{\langle i\rangle} \cup \mathbf{A}_{\langle i\rangle}) \\ &= \wp(\mathbf{O}_{\langle i\rangle} \cup \emptyset)) \\ &= \wp(\{\,\{b\}, \{\}\,\}) \\ &= \{\,\{\}, \{\{b\}, \{\}\}, \{\{b\}\}, \{\{\}\}\,\}\end{aligned}$$

When $m = 2$:

$$\mathbf{A}_{\langle\langle i\rangle\rangle} = \emptyset$$

$$\begin{aligned}\mathbf{A}_{\langle i\rangle} &= \wp(\mathbf{O}_{\langle\langle i\rangle\rangle} \cup \mathbf{A}_{\langle\langle i\rangle\rangle}) \\ &= \wp(\wp(\mathbf{K}_{\langle i\rangle}) \cup \emptyset) \\ &= \wp(\{\ldots\text{all 8 subsets of } \mathbf{K}_{\langle i\rangle}\ldots\}) \\ &= \{\ldots\text{all 256 subsets of } \wp(\mathbf{K}_{\langle i\rangle})\ldots\}\end{aligned}$$

$$\begin{aligned}\mathbf{A}_i &= \wp(\mathbf{O}_{\langle i\rangle} \cup \mathbf{A}_{\langle i\rangle}) \\ &= \wp(\mathbf{O}_{\langle i\rangle} \cup \{\ldots\text{all 256 elements of } \mathbf{A}_{\langle i\rangle}\ldots\}) \\ &= \wp(\{\,\{b\}, \{\}\,\} \cup \{\ldots\text{all 256 elements of } \mathbf{A}_{\langle i\rangle}\ldots\}) \\ &= \wp(\{\{b\}, \ldots\text{all 256 elements of } \mathbf{A}_{\langle i\rangle}\ldots\,\}) \\ &= \wp(\{\ldots\text{all 257 elements of } \mathbf{O}_{\langle i\rangle} \cup \mathbf{A}_{\langle i\rangle}\ldots\}) \\ &= \{\ldots\text{all } 2^{257} \text{ subsets of } \mathbf{O}_{\langle i\rangle} \cup \mathbf{A}_{\langle i\rangle}\ldots\}\end{aligned}$$

But as noted earlier, for the purposes of building a specific model for our analysis of mathematics, we will set the bound of m to 3.

Domains of Quantification and the Proxy Function

The domains over which the variables of our language range are now defined simply as follows. where $\mathbf{D}_t$ is the domain of quantification for type t:

$$\mathbf{D}_t = \mathbf{A}_t \cup \mathbf{O}_t$$

We next define a proxy function $\|\cdot\|$ in two steps. In the first step, we define a *proxy function* $|\cdot|$ so that it maps abstract individuals, abstract properties, and abstract relations to special individuals, special properties, and special relations, respectively. In the second step, we extend this function to the *extended proxy function* $\|\cdot\|$ which is defined on all the domains of quantification: it preserves what $\|\cdot\|$ assigns to the abstract entities but also assigns each ordinary element in each domain of quantification to itself as proxy.

The function $|\cdot|$ is defined generally as, for each type t:

$$|\cdot| : \mathbf{A}_t \to \mathbf{S}_t$$

In the minimal model, there is only one special object in each domain $\mathbf{S}_t$, and so all the abstract objects in $\mathbf{A}_t$ get mapped to the same proxy. For example, in the minimal model:

- where $\mathbf{a}_i$ is an abstract individual in $\mathbf{A}_i$ (i.e., where $\mathbf{a}_i$ is a set of 1st level properties), then $|\mathbf{a}_i| = \boldsymbol{b}$,

- where $\mathbf{a}_{\langle i\rangle}$ is a 1st level abstract property in $\mathbf{A}_{\langle i\rangle}$ (i.e., where $\mathbf{a}_{\langle i\rangle}$ is a set of 2nd level properties of properties), then let $|\mathbf{a}_{\langle i\rangle}| = \boldsymbol{c}$, and

- where $\mathbf{a}_{\langle i,i\rangle}$ is a 1st level abstract relation in $\mathbf{A}_{\langle i,i\rangle}$ (i.e., where $\mathbf{a}_{\langle i,i\rangle}$ is a set of 2nd level properties of relations), then let $|\mathbf{a}_{\langle i,i\rangle}| = \boldsymbol{e}$.

We then extend $|\cdot|$ to the *extended proxy function* $\|\cdot\|$ as follows. Where t is any type and $\mathbf{D}_t$ is the domain of type t (as defined below), and $\boldsymbol{o}$ is a variable ranging over the entities of domain $\mathbf{D}_t$, then for all $\boldsymbol{o} \in \mathbf{D}_t$:

$$\|\boldsymbol{o}\| = \begin{cases} |\boldsymbol{o}| & \text{when } |\boldsymbol{o}| \text{ is defined} \\ \boldsymbol{o} & \text{otherwise} \end{cases}$$

The Model

We now introduce a *bounded domain*, $\mathbf{D}^{n,m}$, as follows:

$$\mathbf{D}^{n,m} =_{df} \{\mathbf{D}_t \mid w(t) \leq n \;\&\; h(t) \leq m\}$$

In other words, a bounded domain collects all the domains of the types t within the width and height bounds n and m.

In order to preserve the information about whether objects in the domain are ordinary or abstract, we define two *indicator* functions $\boldsymbol{A}$ and $\boldsymbol{O}$ defined as follows. Where $\boldsymbol{o}$ is again a variable ranging over the entities of domain $\mathbf{D}_t$, then for all $\boldsymbol{o} \in \mathbf{D}_t$:

$$\boldsymbol{A}(\boldsymbol{o}) = \begin{cases} \mathbf{T} & \text{if } \boldsymbol{o} \in \mathbf{A}_t, \text{ where } \boldsymbol{o} \text{ is of type } t \\ \mathbf{F} & \text{otherwise} \end{cases}$$

$$\boldsymbol{O}(\boldsymbol{o}) = \begin{cases} \mathbf{T} & \text{if } \boldsymbol{o} \in \mathbf{O}_t, \text{ where } \boldsymbol{o} \text{ is of type } t \\ \mathbf{F} & \text{otherwise} \end{cases}$$

We next define an *interpretation* $\mathbf{V}$ that assigns to each constant κ of the language an element of an appropriate domain of quantification, i.e.,

If κ is a constant of type t, $\mathbf{V}(\kappa) \in \mathbf{D}_t$

$\mathbf{V}$ also assigns a special entity to each predicate $A!$ of the language:

$$\mathbf{V}(A!^{\langle t\rangle}) = \mathbf{S}_t$$

That is, the interpretation of the predicate constant $A!^{\langle t\rangle}$ is the set $\mathbf{S}_t$ of proxy elements of type t. Using the definitions above, we then define a model $\mathbf{M}$ as a structure of the form:

$$\mathbf{M} = \langle \mathbf{D}^{n,m}, \|\cdot\|, \boldsymbol{A}, \boldsymbol{O}, \mathbf{V}\rangle$$

where $\mathbf{D}^{n,m}$ is a bounded domain, $\|\cdot\|$ is an extended proxy function, $\boldsymbol{A}$ and $\boldsymbol{O}$ are the indicator functions that identify which elements of the bounded domain are abstract and ordinary, respectively, and $\mathbf{V}$ is an interpretation function. These elements have all been defined as above.

As noted earlier, for analysis of mathematics developed in this paper, we need models with bounded domain $\mathbf{D}^{2,3}$. We leave the complete list of types included within this bound to a footnote.[39] Note that many of these types don't play a role in our analysis of mathematics and aren't represented in Figure 1.

Simultaneous Definition of Denotation and Truth

Assignments to the Variables

In the usual way, an assignment $\boldsymbol{f}$ to the variables is a function that takes each variable in the language to an element of the domain over which the variable ranges. Strictly speaking, $\boldsymbol{f}$ should be relativized to the model $\mathbf{M}$, but we now always suppress the index to $\mathbf{M}$. More specifically:

- If α^t is a variable of type t, $\boldsymbol{f}(\alpha^t) \in \mathbf{D}_t$.

Moreover, where α^t is a variable of type t and $\boldsymbol{o}$ is an object in the domain $\boldsymbol{D}_t$, we use the notation $\boldsymbol{f}[\alpha/\boldsymbol{o}]$ to refer to the assignment function just like $\boldsymbol{f}$ except that it assigns the object $\boldsymbol{o}$ to the variable α. And where $\alpha_1,\ldots,\alpha_n$ are variables of type $t_1,\ldots,t_n$, respectively, and $\boldsymbol{o}_1,\ldots,\boldsymbol{o}_n$ are objects in the domains $\boldsymbol{D}_{t_1},\ldots,\boldsymbol{D}_{t_n}$, respectively, we use the notation

[39] We list the types for each width and height:

- $w=1, h=0$: $\mathbf{D}_i$
- $w=1, h=1$: $\mathbf{D}_{\langle\,\rangle}$, $\mathbf{D}_{\langle i\rangle}$
- $w=1, h=2$: $\mathbf{D}_{\langle\langle\,\rangle\rangle}$, $\mathbf{D}_{\langle\langle i\rangle\rangle}$
- $w=1, h=3$: $\mathbf{D}_{\langle\langle\langle\,\rangle\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle i\rangle\rangle\rangle}$
- $w=2, h=0$: (empty)
- $w=2, h=1$: $\mathbf{D}_{\langle i,i\rangle}$
- $w=2, h=2$: $\mathbf{D}_{\langle i,\langle\rangle\rangle}$, $\mathbf{D}_{\langle i,\langle i\rangle\rangle}$, $\mathbf{D}_{\langle\langle\rangle,i\rangle}$, $\mathbf{D}_{\langle\langle i\rangle,i\rangle}$, $\mathbf{D}_{\langle\langle\rangle,\langle\rangle\rangle}$, $\mathbf{D}_{\langle\langle\rangle,\langle i\rangle\rangle}$, $\mathbf{D}_{\langle\langle i\rangle,\langle\rangle\rangle}$, $\mathbf{D}_{\langle\langle i\rangle,\langle i\rangle\rangle}$
- $w=2, h=3$: $\mathbf{D}_{\langle\langle i,\langle\rangle\rangle\rangle}$, $\mathbf{D}_{\langle\langle i,\langle i\rangle\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle\rangle,i\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle i\rangle,i\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle\rangle,\langle\rangle\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle\rangle,\langle i\rangle\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle i\rangle,\langle\rangle\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle i\rangle,\langle i\rangle\rangle\rangle}$, $\mathbf{D}_{\langle i,\langle\langle\rangle\rangle\rangle}$, $\mathbf{D}_{\langle i,\langle\langle i\rangle\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle\rangle\rangle,i\rangle}$, $\mathbf{D}_{\langle\langle\langle i\rangle\rangle,i\rangle}$, $\mathbf{D}_{\langle\langle\rangle,\langle\langle\rangle\rangle\rangle}$, $\mathbf{D}_{\langle\langle\rangle,\langle\langle i\rangle\rangle\rangle}$, $\mathbf{D}_{\langle\langle i\rangle,\langle\langle\rangle\rangle\rangle}$, $\mathbf{D}_{\langle\langle i\rangle,\langle\langle i\rangle\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle\rangle\rangle,\langle\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle i\rangle\rangle,\langle\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle\rangle\rangle,\langle i\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle i\rangle\rangle,\langle i\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle\rangle\rangle,\langle\langle\rangle\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle\rangle\rangle,\langle\langle i\rangle\rangle\rangle}$, $\mathbf{D}_{\langle\langle\langle i\rangle\rangle,\langle\langle\rangle\rangle\rangle}$

$\boldsymbol{f}[\alpha_i/\boldsymbol{o}_i]_{i=1}^n$ to refer to the assignment just like $\boldsymbol{f}$ except that it assigns $\boldsymbol{o}_1, \ldots, \boldsymbol{o}_n$ to $\alpha_1, \ldots, \alpha_n$, respectively.

Denotation and Satisfaction

Relative to the model $\mathbf{M}$ and variable assignment $\boldsymbol{f}$, we next define, by simultaneous recursion, (a) the *denotation* $\boldsymbol{d_f}(\tau)$ of term τ, and (b) $\boldsymbol{f}$ *satisfies* φ. Strictly speaking, $\boldsymbol{d_f}(\tau)$ should also be indexed to the model $\mathbf{M}$, but we now always suppress the index to $\mathbf{M}$:

D1 Where κ is any constant of type t, $\boldsymbol{d_f}(\kappa) = \mathbf{V}(\kappa)$.

D2 Where α is an variable of type t, $\boldsymbol{d_f}(\alpha) = \boldsymbol{f}(\alpha)$.

S1 If Π is a term of type $\langle t_1, \ldots, t_n\rangle$ $(n \geq 1)$, and $\tau_1, \ldots, \tau_n$ any terms of types $t_1, \ldots, t_n$, respectively, then $\boldsymbol{f}$ satisfies $\Pi\tau_1 \ldots \tau_n$ iff (a) $\boldsymbol{d_f}(\tau_1), \ldots, \boldsymbol{d_f}(\tau_n)$ are all defined, (b) $\boldsymbol{O}(\boldsymbol{d_f}(\Pi)) = \mathbf{T}$ (which implies $\boldsymbol{d_f}(\Pi)$ is defined as well), and (c) $\langle \|\boldsymbol{d_f}(\tau_1)\|, \ldots, \|\boldsymbol{d_f}(\tau_n)\|\rangle \in \boldsymbol{d_f}(\Pi)$.

S2 If Π is any constant or variable of type $\langle\,\rangle$, then $\boldsymbol{f}$ satisfies Π iff $\boldsymbol{d_f}(\Pi) = \mathbf{T}$.

S3 If Π is any term of type $\langle t\rangle$ and τ is any term of type t, then $\boldsymbol{f}$ satisfies $\tau\Pi$ iff (a) $\boldsymbol{d_f}(\tau)$ is defined, (b) $\boldsymbol{A}(\boldsymbol{d_f}(\tau)) = \mathbf{T}$, and (c) $\boldsymbol{d_f}(\Pi) \in \boldsymbol{d_f}(\tau)$.

S4 And so on for the clauses for negation, conditionals, universal quantification, etc. E.g., $\boldsymbol{f}$ satisfies $\forall\alpha\varphi$ iff $\forall\boldsymbol{o}(f[\alpha/\boldsymbol{o}]$ satisfies $\varphi)$.

D3 Where $[\lambda\alpha_1 \ldots \alpha_n\, \varphi]$ is any λ-expression $(n \geq 1)$, $\alpha_1, \ldots, \alpha_n$ are variables of type $t_1, \ldots, t_n$, respectively, and $\boldsymbol{o}_1, \ldots, \boldsymbol{o}_n$ are objects of type $t_1, \ldots, t_n$, respectively, then

$$\boldsymbol{d_f}([\lambda\alpha_1 \ldots \alpha_n\, \varphi]) = \{\langle\|\boldsymbol{o}_1\|, \ldots, \|\boldsymbol{o}_n\|\rangle \mid \boldsymbol{f}[\alpha_i/\boldsymbol{o}_i]_{i=1}^n \text{ satisfies } \varphi)\},$$

where this set of n-tuples is an element of $\mathbf{O}_{\langle t_1,\ldots,t_n\rangle}$.[40]

In the special case where $[\lambda\alpha_1 \ldots \alpha_n\, \varphi]$ is *elementary*, i.e., has the form $[\lambda\alpha_1 \ldots \alpha_n\, \Pi\alpha_1 \ldots \alpha_n]$, then the above definition has the consequence that

- if $\boldsymbol{O}(\boldsymbol{d_f}(\Pi)) = \mathbf{T}$, then $\boldsymbol{d_f}([\lambda\alpha_1 \ldots \alpha_n\, \varphi]) = \boldsymbol{d_f}(\Pi)$, and
- if $\boldsymbol{A}(\boldsymbol{d_f}(\Pi)) = \mathbf{T}$, then $\boldsymbol{d_f}([\lambda\alpha_1 \ldots \alpha_n\, \varphi]) = \{\}$ in $\mathbf{O}_{\langle t_1,\ldots,t_n\rangle}$[41]

D4 Where α is any variable of type t and $\boldsymbol{o}$ is any object of type t, then

$$\boldsymbol{d_f}(\iota\alpha\varphi) = \begin{cases} \boldsymbol{o} \in \mathbf{D}_t, \text{if } \boldsymbol{f}[\alpha/\boldsymbol{o}] \text{ satisfies } \varphi\ \&\\ \quad \forall\boldsymbol{o}'(\boldsymbol{f}[\alpha/\boldsymbol{o}'] \text{ satisfies } \varphi \rightarrow \boldsymbol{o}' = \boldsymbol{o})\\ \text{undefined, otherwise} \end{cases}$$

D5 And so on for the other cases where φ is a complex term of type $\langle\,\rangle$, i.e., where φ is any complex propositional formula. E.g., $\boldsymbol{d_f}(\neg\varphi) = \mathbf{T}$ iff $\boldsymbol{d_f}(\varphi) = \mathbf{F}$.

Truth

In the usual manner we say that φ is true just in case every assignment $\boldsymbol{f}$ satisfies φ.

Axioms

Since we've assumed classical logic in the description of our model, it remains to show that the axiom groups of Section 3.2, 3.3, and 3.4 are true in the above model. It is easy to see that the following lemma holds:

> **Substitution Lemma**. If $\boldsymbol{d_f}(\tau)$ is defined and φ is any formula, then $\boldsymbol{f}$ satisfies φ^τ_α if and only if $\boldsymbol{f}[\alpha/\boldsymbol{d_f}(\tau)]$ satisfies φ.

In other words, $\boldsymbol{f}$ satisfies φ^τ_α whenever the assignment just like $\boldsymbol{f}$ except that it assigns $\boldsymbol{d_f}(\tau)$ to α satisfies φ (assuming $\boldsymbol{d_f}(\tau)$ is defined). This Lemma holds because, according to our semantics, the truth value of an atomic formula is calculated in terms of the denotations of its terms, assuming they all have such, and this feature is inherited by all the molecular and quantified formulas built out of such formulas.

In what follows, we omit the proofs that the axioms for classical logic hold in our model. We also omit proofs for the Hintikka (1959) axiom for descriptions (AXIOM 1) and AXIOMS 4 – 7. These are obviously true in the model (since the model was constructed in part to make these axioms

[40] In the system we've developed here, every λ-expression gets a denotation. However, as noted above, in more recent formulations of object theory, λ-expressions with encoding formulas are allowed, and a free logic governs them, to allow for non-denoting paradoxical predicates, such as the Kirchner predicate (in these paradoxical predicates, the variable bound by the λ-bound occurs within a non-propositional matrix.

[41] In other words, if the head relation term in an elementary λ-expression denotes an abstract relation, then the λ-expression denotes an *ordinary* relation (of type $\langle t_1, \ldots, t_n\rangle$ that is never exemplified.

true).[42] It remains to show only that the three distinctive principles, the AXIOM for the substitution of identicals, and AXIOMS 2 and 3, are true in the model.

Axiom: Substitution of Identicals

The proof that substitution of identicals is true in the model is by cases. The cases are:

$x = y$
$F = G$
$R = S$
$p = q$

Consider the first case:

> Assume some assignment, say $\boldsymbol{f}$, satisfies $x = y$. We want to show $\boldsymbol{f}$ satisfies φ if and only if $\boldsymbol{f}$ satisfies φ', where φ' is the result of replacing 0 or more free occurrences of x by y in φ. Now our assumption implies, by definition of $x = y$, that $\boldsymbol{f}$ satisfies:
>
> $$(O!x \,\&\, O!y \,\&\, \forall F(Fx \equiv Fy)) \vee (A!x \,\&\, A!y \,\&\, \forall F(xF \equiv yF))$$
>
> But since there are no ordinary individuals in the model (the domain $\mathbf{O}_i$ is empty), it follows that $\boldsymbol{f}$ satisfies:
>
> $$A!x \,\&\, A!y \,\&\, \forall F(xF \equiv yF)$$
>
> So by S4, $\boldsymbol{f}$ satisfies $\forall F(xF \equiv yF)$. Now by S3, we know:
>
> $\boldsymbol{f}$ satisfies 'xF' iff (a) $\boldsymbol{d_f}(x)$ is defined, (b) $\boldsymbol{A}(\boldsymbol{d_f}(x)) = \mathbf{T}$, and (c) $\boldsymbol{d_f}(F) \in \boldsymbol{d_f}(x)$.
>
> And we know something analogous for $\boldsymbol{f}$ satisfies yF. Now suppose for reductio that $\boldsymbol{d_f}(x) \neq \boldsymbol{d_f}(y)$. Since both objects are in $\mathbf{A}_i$, they are both sets of type $\langle i \rangle$ properties, so there must be a property that is an element of one, say, $\boldsymbol{d_f}(x)$, and not the other, that is, $\boldsymbol{d_f}(y)$. Let κ be such a property, so that we know $\kappa \in \boldsymbol{d_f}(x)$ and not $\kappa \in \boldsymbol{d_f}(y)$. Now consider the variable assignment $\boldsymbol{f}[F/\kappa]$. Since $\kappa \in \boldsymbol{d_f}(x)$, it follows by S3 that $\boldsymbol{f}[F/\kappa]$ satisfies xF. And since $\kappa \notin \boldsymbol{d_f}(y)$, it follows by S3 that $\boldsymbol{f}[F/\kappa]$ doesn't satisfy yF. So by the biconditional clause of S4, $\boldsymbol{f}[F/\kappa]$ doesn't satisfy $xF \equiv yF$. Then, by the universal quantifier clause of S4, $\boldsymbol{f}$ doesn't satisfy $\forall F(xF \equiv yF)$, This contradicts the assumed fact that $\boldsymbol{f}$ does satisfies $\forall F(xF \equiv yF)$. Hence, $\boldsymbol{d_f}(x) = \boldsymbol{d_f}(y)$. So by reasoning from the Substitution Lemma, we can argue as follows: $\boldsymbol{f}$ satisfies φ if and only if $\boldsymbol{f}[x/\boldsymbol{d_f}(x)]$ satisfies φ iff $\boldsymbol{f}[x/\boldsymbol{d_f}(y)]$ satisfies φ iff $\boldsymbol{f}$ satisfies φ'.

Given the definitions for identity in Section 2.2, the proof of the remaining cases, i.e., $F = G$, $R = S$, and $p = q$, are similar to the above.[43]

AXIOM 2: β-Conversion

To see AXIOM 2 holds, note that the axiom is trivially true in the case where $[\lambda\alpha_1 \ldots \alpha_n\, \varphi]$ is an elementary λ-expression and φ is any propositional formula. For in that case, $[\lambda\alpha_1 \ldots \alpha_n\, \varphi]$ has the form:

$$[\lambda\alpha_1 \ldots \alpha_n\, \Pi\alpha_1 \ldots \alpha_n],$$

where $\alpha_1, \ldots, \alpha_n$ have types $\langle t_1, \ldots, t_n \rangle$, respectively, and Π is any n-place relation term of type $\langle t_1, \ldots, t_n \rangle$. So $[\lambda\alpha_1 \ldots \alpha_n\, \Pi\alpha_1 \ldots \alpha_n]\beta_1 \ldots \beta_n \equiv \Pi\beta_1 \ldots \beta_n$ holds, by the consequence noted at the end of D3. For there are two cases: if Π denotes an ordinary relation, then the denotation of the λ-expression is just the denotation of Π itself, and if Π denotes an abstract relation, then Π denotes the empty set and both sides of the biconditional $[\lambda\alpha_1 \ldots \alpha_n\, \Pi\alpha_1 \ldots \alpha_n]\beta_1 \ldots \beta_n \equiv \Pi\beta_1 \ldots \beta_n$ are false (given that the contrapositive of AXIOM 5 tells us that abstract relations aren't exemplified).

When $[\lambda\alpha_1 \ldots \alpha_n\, \varphi]$ is non-elementary and φ is any propositional formula, our semantic definition of truth requires us to show that every assignment function $\boldsymbol{f}$ satisfies:

[42] For example, AXIOM 4 asserts that if an object is distinct from the null object, then it is abstract if and only if it encodes a property. AXIOM 5 asserts that if a relation is exemplified, it is not abstract. AXIOM 6 asserts that λ-expressions don't denote abstract relations. And AXIOM 7 asserts that η-Conversion holds for elementary λ-expressions in which the head relation is ordinary. It should be relatively straightforward to see that these are all true in the model.

[43] For example, if $\boldsymbol{f}$ satisfies $F = G$, it satisfies $(O!F \,\&\, O!G \,\&\, \forall x(xF \equiv xG)) \vee (A!F \,\&\, A!G \,\&\, \forall \mathcal{H}(F\mathcal{H} \equiv G\mathcal{H}))$. If the values of F and G are both in $\mathbf{O}_{\langle i \rangle}$ and the same individuals encode them (which in the model means they are members of the same members of $\mathbf{A}_i$), then they are identical. And if the values of F and G are both in $\mathbf{A}_{\langle i \rangle}$, and they encode the same properties of properties (which in the model means they have the same members), they are identical.

$[\lambda\alpha_1 \ldots \alpha_n\, \varphi]\beta_1 \ldots \beta_n \equiv \varphi^{\beta_1 \ldots \beta_n}_{\alpha_1,\ldots,\alpha_n}$, provided β_i is substitutable for α_i in φ $(1 \leq i \leq n)$

For simplicity and ease of readability, we prove this only for the 1-place case, so that where α, β are variables of some type t, our task is to show that every assignment function $\boldsymbol{f}$ satisfies:

$[\lambda\alpha\, \varphi]\beta \equiv \varphi^{\beta}_{\alpha}$, provided β is substitutable for α in φ

Moreover, if we use $x, y, z, \ldots$ as arbitrarily chosen variables of type t, we simply have to show:

$[\lambda x\, \varphi]y \equiv \varphi^{y}_{x}$, provided y is substitutable for x in φ

So suppose $\boldsymbol{f}$ is an arbitrary assignment function. Then, by the clauses in S4, we have to show $\boldsymbol{f}$ satisfies $[\lambda x\, \varphi]y$ if and only if $\boldsymbol{f}$ satisfies φ^{y}_{x}:

($\rightarrow$) Assume $\boldsymbol{f}$ satisfies $[\lambda x\, \varphi]y$, to show $\boldsymbol{f}$ satisfies φ^{y}_{x}. Then by clause S1, our assumption implies:

(a) $\boldsymbol{d}_f(y)$ is defined

(b) $\boldsymbol{O}(\boldsymbol{d}_f([\lambda x\, \varphi])) = \mathbf{T}$, i.e., $\boldsymbol{d}_f([\lambda x\, \varphi]) \in \mathbf{O}_{\langle t\rangle}$, by definition of $\boldsymbol{O}$

(c) $\|\boldsymbol{d}_f(y)\| \in \boldsymbol{d}_f([\lambda x\, \varphi])$

From (c) and clause D3, it follows that:

$\|\boldsymbol{d}_f(y)\| \in \{\|\boldsymbol{o}\| \mid \boldsymbol{f}[x/\boldsymbol{o}] \text{ satisfies } \varphi)\}$

i.e., by set-abstraction, the fact that φ is propositional, and the **Lemma on Proxies** (see below):

$\boldsymbol{f}[x/\boldsymbol{d}_f(y)]$ satisfies φ

i.e., by the Substitution Lemma,

$\boldsymbol{f}$ satisfies φ^{y}_{x}

For this conclusion to hold, it remains only to show:

Lemma on Proxies: Let x be a variable of type t. Then if φ is propositional and $\boldsymbol{o}$ and $\boldsymbol{o}'$ have the same proxy, then:

L1 for any term τ in φ, $\|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau)\| = \|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau)\|$, and

L2 $f[x/\boldsymbol{o}]$ satisfies φ iff $f[x/\boldsymbol{o}']$ satisfies φ

Proof. The proof proceeds by induction on the λ-rank of propositional formulas φ, i.e., how deeply nested is the deepest λ expressions in φ. But in what follows, the notion of λ-rank applies to any formula or term.[44] (We ignore the trivial case where x doesn't occur free in φ.) For L2, without loss of generality, we need only prove the left to right direction, i.e., that if $f[x/\boldsymbol{o}]$ satisfies φ, then $f[x/\boldsymbol{o}']$ satisfies φ. So assume $f[x/\boldsymbol{o}]$ satisfies φ.

The base case is where φ is a formula of λ-rank 0, i.e., without λ-expressions. We first consider atomic formulas φ of the form $\Pi^n\tau_1 \ldots \tau_n$ $(n \geq 0)$ where none of Π^n, $\tau_1, \ldots, \tau_n$ are λ-expressions. Note that to prove L1, we must show that (1) $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\Pi) = \boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi)$ and (2) for each τ_i, $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau_i) = \boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau_i)$. We will prove (1) and (2) in the course of proving L2. There are two subcases for L2: $n \geq 1$ or $n = 0$. When $n \geq 1$, then each of Π^n, $\tau_1, \ldots, \tau_n$ is either a constant, a variable, or a description. And when $n = 0$, then φ has the form Π where Π is of the empty type, and Π can only be a constant or a variable, since there are no descriptions of the empty type. We cover these two subcases in turn.

In the first subcase $(n \geq 1)$, φ is governed by S1. So φ has the form $\Pi^n\tau_1 \ldots \tau_n$ and contains no λ-expressions. So we know[45] (i) x is either Π^n or one of the τ_i, (ii) $\tau_1, \ldots, \tau_n$ have any types and (iii) Π is of an appropriate type to relate $\tau_1, \ldots, \tau_n$. Then, since $f[x/\boldsymbol{o}]$ satisfies φ, it follows that:

(a) $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau_1), \ldots, \boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau_n)$ are all defined

(b) $\boldsymbol{O}(\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\Pi)) = \mathbf{T}$

(c) $\langle\|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau_1)\|, \ldots, \|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau_n)\|\rangle \in \boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\Pi)$.

We're trying to show $\boldsymbol{f}[x/\boldsymbol{o}']$ satisfies φ, i.e., all of (d) – (f) have to hold:

(d) $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau_1), \ldots, \boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau_n)$ are all defined.

(e) $\boldsymbol{O}(\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi)) = \mathbf{T}$.

[44] If φ (or τ) contains no λ-expressions, it has a λ-rank of 0. If no λ-expression in φ (τ) contains a λ-expression, then its λ-rank is 1. If φ (τ) contains a λ-expression whose matrix has λ-rank n and no λ-expression in φ (τ) has a λ-rank greater than n, then φ (τ) has a λ-rank of $n+1$.

[45] We need not consider the case where Π contains x free as proper subterm because the only way for that to happen is if Π is a λ-expression (which is ruled in this base case), or Π is description (which is ruled out because descriptions can't have free variables).

(f) $\langle \|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau_1)\|, \ldots, \|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau_n)\|\rangle \in \boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi)$.

Proof of (d). This follows from (a) because for any term τ, if $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau)$ is defined, then $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau)$ is defined. (If $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau)$ were undefined, then τ has to be a description. But there are only closed descriptions in the language so its denotation is independent of any particular assignment to the variables. So by D4, since $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau)$ is defined, the description has a denotation, and so it must have a denotation for any assignment to the variables. So $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau)$ is defined. Contradiction.)

Proof of (e). This follows from (b) by cases. (i) If Π is a constant or a variable other than x, then $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\Pi) = \boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi)$, because $\boldsymbol{f}[x/\boldsymbol{o}]$ and $\boldsymbol{f}[x/\boldsymbol{o}']$ differ only by their assignment to the variable x which is different than Π. So if $\boldsymbol{O}(\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\Pi)) = \mathbf{T}$, then $\boldsymbol{O}(\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi)) = \mathbf{T}$. (ii) If Π is x, then by (b), $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\Pi)) = \boldsymbol{o}$. Since $\boldsymbol{o}$ and $\boldsymbol{o}'$ have the same proxy and $\boldsymbol{O}(\boldsymbol{o}) = \mathbf{T}$ (i.e., $\boldsymbol{o}$ is ordinary), $\boldsymbol{o} = \boldsymbol{o}'$. So $\boldsymbol{O}(\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi)) = \mathbf{T}$ (iii) If Π is a description, then x doesn't occur free and so $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\Pi)) = \boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi))$ and it follows that $\boldsymbol{O}(\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi)) = \mathbf{T}$. Note that we have now proved part (1) of L1 where φ falls under the first subcase.

Proof of (f). There are 2 cases to consider: one or more of the τ_i is x or Π is x. Suppose one or more of the τ_i is x. Then we note that $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\Pi) = \boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi)$ because x does not occur in Π (because Π isn't a λ-expression and x has to have a different type from Π and so can't be Π). Moreover, for each τ_i, $\|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau_i)\| = \|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau_i)\|$. If τ_i is x in which case this follows from the fact that $\boldsymbol{o}$ and $\boldsymbol{o}'$ have the same proxy. Otherwise, τ_i is a constant, a description, or a variable other than x and so the denotation of τ_i under $\boldsymbol{f}[x/\boldsymbol{o}']$ is the same as that under $\boldsymbol{f}[x/\boldsymbol{o}]$. The claim then follows from (c). Alternatively, suppose then that Π is x. Then $\|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau_i)\| = \|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau_i)\|$ (since in this case, x isn't one of the τ_i and can't occur free in any τ_i that is a definite description). Since, by (b), Π is ordinary, then by the argument in (e), $\boldsymbol{o} = \boldsymbol{o}'$, and so the claim is trivially true by (c). Note that we have now proved part (2) of L1 where φ falls under the first subcase.

In the second subcase ($n = 0$), φ has the form Π where Π is of the empty type. So Π can only be a constant or a variable, since there are no descriptions of the empty type. If φ is a constant or variable of type $\langle\,\rangle$, S2 applies. Then either Π is either a constant, or a variable other than x, or x itself. If Π is a constant or a variable other than x, then $\boldsymbol{f}[x/\boldsymbol{o}]$ satisfies Π iff $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\Pi) = \mathbf{T}$ (by S2) iff $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi) = \mathbf{T}$ (since Π doesn't contain a free occurrence of x) iff $\boldsymbol{f}[x/\boldsymbol{o}']$ satisfies Π (by S2). If Π is x, then $\boldsymbol{f}[x/\boldsymbol{o}]$ satisfies φ iff $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(x) = \mathbf{T}$ (by S2) iff $\boldsymbol{o} = \mathbf{T}$ iff $\boldsymbol{o}' = \mathbf{T}$ (see below) iff $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(x) = \mathbf{T}$ iff $\boldsymbol{f}[x/\boldsymbol{o}']$ satisfies φ. To see that $\boldsymbol{o} = \mathbf{T}$ iff $\boldsymbol{o}' = \mathbf{T}$, recall that $\|\boldsymbol{o}'\| = \|\boldsymbol{o}\|$ (by hypothesis) and since $\mathbf{T} \notin \mathbf{S}_{\langle\,\rangle}$, the only object with $\mathbf{T}$ at its proxy is $\mathbf{T}$ itself. So $\|\boldsymbol{o}'\| = \|\boldsymbol{o}\| = \mathbf{T}$ iff $\boldsymbol{o}' = \boldsymbol{o} = \mathbf{T}$. Note that this proves part (1) of L1 where φ falls under the second subcase. There is no part (2) of L1 for this subcase.

We now have that L1 and L2 hold for atomic formulas φ of λ-rank 0. We conclude this base case by noting that L1 and L2 hold for complex formulas φ of the form $\neg\psi$, $\psi \to \chi$, and $\forall\alpha\psi$. By S4 and D5, the truth of L1 and L2 for these complex formulas is grounded in the truth of L1 and L2 for the atomic formulas with no λ-expressions given that complex formulas of rank 0 contain no λ-expressions.

Inductive cases: IH: The lemma holds for ψ with λ-rank of n or less, i.e., for ψ with λ-rank of n or less we may assume:

IH-L1: $\|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau)\| = \|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau)\|$, for any term τ in ψ, and

IH-L2: $f[x/\boldsymbol{o}]$ satisfies ψ if and only if $f[x/\boldsymbol{o}']$ satisfies ψ.

We need to show that it holds for φ with λ-rank of $n+1$.

To show L1, we need to show that $\|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau)\| = \|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau)\|$, for any term τ in φ. We first consider the case where φ is atomic. Fix an arbitrary such τ and consider its denotation:

- D1 and D2 only apply if τ has a λ-rank of 0, so the result follows immediately from IH-L1.

- D3 applies when τ is of the form $[\lambda\alpha_1 \ldots \alpha_n\,\psi]$. Now if τ has λ-rank n or less, then the result follows by IH-L1. So we need only be concerned with the case when τ has λ-rank $n+1$ where ψ has λ-rank n. We have to show $\|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}([\lambda\alpha_1 \ldots \alpha_n\psi])\| = \|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}([\lambda\alpha_1 \ldots \alpha_n\psi])\|$. Note that D3 defines the denotation of the λ-expression as the set of n-tuples (of proxies of objects) such that $\boldsymbol{f}[\alpha_i/\boldsymbol{o}_i]_{i=1}^n$ satisfies ψ. But ψ has λ-rank n (one less than τ), so by IH-L2, $\boldsymbol{f}[\alpha_i/\boldsymbol{o}_i]_{i=1}^n$ satisfies ψ if and only if $\boldsymbol{f}[\alpha_i/\boldsymbol{o}'_i]_{i=1}^n$ satisfies ψ, so the set of proxy-tuples must be the same and L1 holds.

- D4 applies if τ is a description. But the result holds trivially since x can't be free in τ.

We now have the L1 holds for atomic formulas φ of λ-rank up to $n+1$. We can then conclude that L1 holds for complex formulas φ of the form $\neg\psi$, $\psi \to \chi$, and $\forall\alpha\psi$. By D5, the truth of L1 for these complex formulas is grounded in the truth of L1 for the atomic formulas with λ-rank up to $n+1$ given that a complex formula can not have a λ-rank higher than its component atomic formulas.

To show L2, we have to show: $f[x/\boldsymbol{o}]$ satisfies φ if and only if $f[x/\boldsymbol{o}']$ satisfies φ, when φ has an λ-rank of $n+1$. We now make use of L1, which has been proved for all λ-ranks. So we know that for any term τ in φ, $\|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau)\| = \|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau)\|$. We first consider the case where φ is atomic. Moreover, we need only consider formulas governed by S1, for (a) if the λ-rank of φ is some n other than 0, then φ is not a constant or variable, and so S2 doesn't apply, (b) S3 doesn't apply since that governs encoding formulas, which aren't propositional. So we are consider φ of the form $\Pi\tau_1\ldots\tau_n$ where any of Π, $\tau_1,\ldots,\tau_n$ can have a λ-rank of $n+1$. Without loss of generality, we need only prove the left to right direction, i.e., that if $f[x/\boldsymbol{o}]$ satisfies φ, then $f[x/\boldsymbol{o}']$ satisfies φ. So assume $f[x/\boldsymbol{o}]$ satisfies φ. By assumption, it follows that:

(a) $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau_1),\ldots,\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau_n)$ are all defined.

(b) $\boldsymbol{O}(\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\Pi)) = \mathbf{T}$.

(c) $\langle\|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau_1)\|,\ldots,\|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau_n)\|\rangle \in \boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\Pi)$.

We're trying to show $\boldsymbol{f}[x/\boldsymbol{o}']$ satisfies φ, i.e., all of (d) – (f) have to hold:

(d) $\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau_1),\ldots,\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau_n)$ are all defined.

(e) $\boldsymbol{O}(\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi)) = \mathbf{T}$.

(f) $\langle\|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau_1)\|,\ldots,\|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau_n)\|\rangle \in \boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi)$.

Proof of (d). This follows by the exact reasoning for clause (d) in the base case, which did not depend on the λ-rank of τ_i.

Proof of (e). There are four cases to consider, 3 of which are identical to the proof of clause (e) in the base case. The 4 cases are (i) Π is a constant or a variable other than x, (ii) Π is x, (iii) Π is a description, or (iv) Π is a λ-expression. So it remains to show only (iv). Π is a λ-expression and by D3, the denotation of all λ-expressions are ordinary. Thus (e) follows since $\boldsymbol{O}(\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi)) = \mathbf{T}$.

Proof of (f). By assumption (b), we know that the denotation of Π is ordinary. And since the denotation of Π is ordinary, the proxy of the denotation is just equal to the denotation itself. So:

$$\begin{aligned}
\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\Pi) &= \|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\Pi)\| && \text{(since Π is ordinary)}\\
&= \|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi)\| && \text{(by L1)}\\
&= \boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\Pi) && \text{(since Π is ordinary)}
\end{aligned}$$

Moreover, we also know, by L1, that the proxies of the denotations for each τ are identical, that is for each τ_i $\|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}]}(\tau_i)\| = \|\boldsymbol{d}_{\boldsymbol{f}[x/\boldsymbol{o}']}(\tau_i)\|$. So the tuple in (f) is identical to the tuple in (c). Hence it follows that (f) must hold if (c) holds.

We now have that L2 holds for atomic formulas φ of λ-rank $n+1$. We conclude this inductive case by noting that L2 holds for complex formulas φ of the form $\neg\psi$, $\psi \to \chi$, and $\forall\alpha\psi$. By S4, truth of L2 for these complex formulas is grounded in the truth of L2 for the atomic formulas with λ-rank less than or equal to $n+1$, given that complex formulas only have rank $n+1$ in virtue of the ranks of their atomic components. That is, S4 doesn't increase the λ-rank.

AXIOM 3: Comprehension Principles for Abstracta

At the beginning of the appendix, we noted that our analysis of mathematics does not require the full unbounded language. In general we need only validate comprehension up to a given fixed type height h. So we fix n and m such that the bounded language $\mathcal{L}_{n,m}$ has $m > h$. Now we will show AXIOM 3 holds in the model for any arbitrary t such that $h(t) < m$. Recall that by definition (2) of $\mathbf{A}_t$ (in the Appendix subsection Structural Domains: Abstract Objects), when $h(t) < m$, we have that $\mathbf{A}_t = \wp(\mathbf{O}_{\langle t\rangle} \cup \mathbf{A}_{\langle t\rangle})$. We now show that $\exists\alpha(A!\alpha \,\&\, \forall F(\alpha F \equiv \varphi))$ holds by showing that $\{F \,|\, \varphi\}$ defines an abstract object of type t.

Proof. Pick an arbitrary formula φ, where α of type t doesn't occur free in φ. To show that the instance $\exists\alpha(A!\alpha \& \forall F(\alpha F \equiv \varphi))$ is true in the model, we have to show that every assignment $\boldsymbol{f}$ satisfies this instance. So pick an arbitrary assignment $\boldsymbol{f}$. By S4 (and the definition of the existential quantifier), we have to show that for some element, say $\boldsymbol{o}^*$, in the domain over which α ranges, $\boldsymbol{f}[\alpha/\boldsymbol{o}^*]$ satisfies $A!\alpha \,\&\, \forall F(\alpha F \equiv \varphi)$. So again by S4, we have to show that for some element $\boldsymbol{o}^*$ in the domain over which α ranges, (a) $\boldsymbol{f}[\alpha/\boldsymbol{o}^*]$ satisfies $A!\alpha$ and (b) $\boldsymbol{f}[\alpha/\boldsymbol{o}^*]$ satisfies $\forall F(\alpha F \equiv \varphi)$.

So we show that when we choose the set $\{\boldsymbol{o} \mid \boldsymbol{f}[F/\boldsymbol{o}] \text{ satisfies } \varphi\}$ as our witnesss, both (a) and (b) hold.

(a) To show:

$$\boldsymbol{f}[\alpha/\{\boldsymbol{o} \mid \boldsymbol{f}[F/\boldsymbol{o}] \text{ satisfies } \varphi\}] \text{ satisfies } A!\alpha$$

we have to show, by S1, that:

(i) $\boldsymbol{d}_{\boldsymbol{f}[\alpha/\{\boldsymbol{o}|\boldsymbol{f}[F/\boldsymbol{o}] \text{ satisfies } \varphi\}]}(\alpha)$ is defined,

(ii) $\boldsymbol{O}(\boldsymbol{d}_{\boldsymbol{f}[\alpha/\{\boldsymbol{o}|\boldsymbol{f}[F/\boldsymbol{o}] \text{ satisfies } \varphi\}]}(A!)) = \mathbf{T}$, and

(iii) $\|\boldsymbol{d}_{\boldsymbol{f}[\alpha/\{\boldsymbol{o}|\boldsymbol{f}[F/\boldsymbol{o}] \text{ satisfies } \varphi\}]}(\alpha)\| \in \boldsymbol{d}_{\boldsymbol{f}[\alpha/\{\boldsymbol{o}|\boldsymbol{f}[F/\boldsymbol{o}] \text{ satisfies } \varphi\}]}(A!)$.

I.e., by applying definitions, we have to show:

(i) $\{\boldsymbol{o} \mid \boldsymbol{f}[F/\boldsymbol{o}] \text{ satisfies } \varphi\} \in \mathbf{A}_t$.

(ii) $A!$ is ordinary.

(iii) $\|\{\boldsymbol{o} \mid \boldsymbol{f}[F/\boldsymbol{o}] \text{ satisfies } \varphi\}\| \in \mathbf{V}(A!^{\langle t\rangle}) = \mathbf{S}_t$.

(ii) is trivial; (iii) follows directly from (i) by definition of the proxy function, and so it remains to show (i).

Note that $\{\boldsymbol{o} \mid \boldsymbol{f}[F/\boldsymbol{o}] \text{ satisfies } \varphi\}$ is a fixed set of objects of type $\langle t\rangle$. But all objects of type $\langle t\rangle$ are in $\mathbf{O}_{\langle t\rangle}$ or $\mathbf{A}_{\langle t\rangle}$. Since $h(t) < m$, it follows that $\{\boldsymbol{o} \mid \boldsymbol{f}[F/\boldsymbol{o}] \text{ satisfies } \varphi\} \in \wp(\mathbf{O}_{\langle t\rangle} \cup \mathbf{A}_{\langle t\rangle}) = \mathbf{A}_t$.

Bibliography

Anderson, D., and E. Zalta, 2004, 'Frege, Boolos, and Logical Objects', *Journal of Philosophical Logic*, **33**/1 (February): 1–26.

Benacerraf, P., 1981, 'Frege: The Last Logicist', in *Midwest Studies in Philosophy: VI*, P. French *et al.* (eds.), Minneapolis: U. Minnesota Press; page reference is to the reprint in Demopoulos [1995].

Blanchette, P., 2012, *Frege's Conception of Logic*, New York: Oxford University Press.

Boolos, G., 1986, 'Saving Frege From Contradiction', *Proceedings of the Aristotelian Society*, 87: 137–151; reprinted in Boolos [1998], 171–182.

Boolos, G., 1990, 'The Standard of Equality of Numbers,' in *Meaning and Method: Essays in Honor of Hilary Putnam*, G. Boolos (ed.), Cambridge: Cambridge University Press, pp. 261-277; page reference is to the reprint in Boolos [1998], pp. 202–19.

Boolos, G., 1998, *Logic, Logic, and Logic*, Cambridge MA: Harvard University Press.

Brandom, R., 1998, *Making it Explicit*, Cambridge, MA: Harvard University Press.

Bueno, O., C. Menzel, and E. Zalta, 2014, 'Worlds and Propositions Set Free,' *Erkenntnis*, 79: 797–820.

Burgess, J., 2003, 'Review of Kit Fine, *The Limits of Abstraction*', *Notre Dame Journal of Formal Logic*, 44/4: 227–251.

Burgess, J., 2005, *Fixing Frege*, Princeton: Princeton University Press.

Carnap, R., 1931, 'Die logizistische Grundlegung der Mathematik', *Erkenntnis*, 2: 91–105.

Carnap, R., 1947, *Meaning and Necessity*, Chicago: University of Chicago Press.

Church, A., 1951, 'A Formulation of the Logic of Sense and Denotation', in *Structure, Method and Meaning: Essays in Honor of Henry M. Sheffer*, P. Henle *et al.* (eds.), New York: Liberal Arts Press, 3–24.

Cook, R., 2003, 'Iteration One More Time', *Notre Dame Journal of Formal Logic*, 44(2): 63–92.

Dummett, M., 1991, *Frege: Philosophy of Mathematics*, Cambridge, MA: Harvard University Press.

Field, H., 1989, *Realism, Mathematics, and Modality*, Oxford: Blackwell.

Fine, K., 2002, *The Limits of Abstraction*, Oxford: Clarendon Press.

Francez, N., and, R. Dyckhoff, 2010, 'Proof-theoretic Semantics for a Natural Language Fragment', *Linguistics and Philosophy*, 33: 447–477.

Friedman, M., 1994 [1999], "Geometry, Convention, and the Relativized *A Priori*", in W. Salmon and G. Wolters (eds.), *Logic, Language, and the Structure of Scientific Theories*, Pittsburgh: University of Pittsburgh Press, pp. 21–34; reprinted under the title "Geometry, Relativity, and Convention", in M. Friedman, *Reconsidering Logical Positivism*, Cambridge: Cambridge University Press. (Page reference is to the reprint.)

Frege, G., 1893/1901, *Grundgesetze der Arithmetik*, 2 volumes, Band I (1893), Band II (1903), Jena: Verlag Hermann Pohle.

Göedel, K., 1964, 'What is Cantor's Continuum Problem,' reprinted in S. Feferman et al., *Kurt Göedel: Collected Works* (Volume II: Publications 1938–1974), Oxford: Oxford University Press, 1990, 254–270.

Goldfarb, W., 2001, 'Frege's Conception of Logic,' in J. Floyd and S. Shieh (eds.), *Future Pasts: The Analytic Tradition in Twentieth Century Philosophy*, Oxford: Oxford University Press, 25-41.

Hale, B., 1987, *Abstract Objects*, Oxford: Blackwell.

Heck, R., 1992, 'On the consistency of second-order contextual definitions,' *Noûs*, 26: 491–94.

Hintikka, J. 1959, 'Towards a Theory of Definite Descriptions,' *Analysis*, 19(4): 79–85.

Hodes, H., 1984, 'Logicism and the Ontological Commitments of Arithmetic,' *The Journal of Philosophy*, 81(3): 123–149.

Hodes, H., 1991, 'Where Do Sets Come From?', *Journal of Symbolic Logic*, 56(1): 151–175.

Kirchner, D., 2017, *Representation and Partial Automation of the Principia Logico-Metaphysica in Isabelle/HOL*, Master's Thesis, Institut für Mathematik, Freie Universität Berlin.

Klev, A.M., 2017, 'Dedekind's Logicism', *Philosophia Mathematica*, 25(3): 341–36.

Leitgeb, H., and Ladyman, J., 2008, 'Criteria of Identity and Structuralist Ontology', *Philosophia Mathematica* 16(3): 388–396.

Linnebo, O., 2003, 'Frege's conception of logic: From Kant to Grundgesetze,' *Manuscrito - Rev. Int. Fil. Campinas*, 26(2): 235-252.

Linnebo, O., 2009, 'Introduction', in Linnebo (ed.) 2009, pp. 321–329.

Linnebo, O. (ed.), 2009, *The Bad Company Problem*, special issue, *Synthese*, 70/3 (October).

Linsky, B., and Zalta, E., 1995, 'Naturalized Platonism vs Platonized Naturalism', *The Journal of Philosophy*, 92(10): 525–555.

Linsky, B., and Zalta, E., 2006, 'What is Neologicism?', *Bulletin of Symbolic Logic*, 12/1: 60–99.

MacBride, F., 2003, 'Speaking with Shadows: A Study of Neo-Logicism,' *British Journal for the Philosophy of Science*, 54: 103–163.

MacFarlane, J., 2002, 'Frege, Kant, and the Logic in Logicism', *The Philosophical Review*, 111 (1): 25–65.

Montague, R., 1973, 'The proper treatment of quantification in ordinary English', *Approaches to Natural Language*, in K.J.J. Hintikka, J.M.E. Moravcsik, and P. Suppes (eds.), D. Reidel, Dordrecht, 1973, pp. 221–242; reprinted in R. Thomason (ed.), *Formal Philosophy: Selected Papers of Richard Montague*, New Haven: Yale University Press.

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 Zalta, E., forthcoming, 'Foundations for Mathematical Structuralism', *Mind*.

Pelletier, F.J., and Zalta, E., 2000, 'How to Say Goodbye to the Third Man', *Noûs*, 34/2 (June): 165–202.

Prawitz, D., 2006, 'Meaning Approached via Proofs', *Synthese*, 148: 507–524.

Rayo, A., 2005, 'Logicism Reconsidered', in Stewart Shapiro (ed.), *The Oxford Handbook of Philosophy of Mathematics and Logic*, New York: Oxford University Press, 203–235; doi:10.1093/0195148770.003.0007

Roeper, P., 2015, 'A Vindication of Logicism', *Philosophia Mathematica*, 24(3): 360–378.

Schroeder-Heister, P., 2006, 'Validity Concepts in Proof-Theoretic Semantics', *Synthese*, 148: 525–571.

Sellars, W., 1980, 'Inference and Meaning', in *Pure Pragmatics and Possible Worlds*, J. Sicha (ed.), Atascadero: Ridgeview, pp. 261–313.

Tarski, A., 1933, 'Pojęcie prawdy w językach nauk dedukcyjnych', *Prace Towarzystwa Naukowego Warszawskiego, Wydzial III Nauk Matematyczno-Fizycznych*, 34, Warsaw; translated as 'The Concept of Truth in Formalized Languages',' by J.H. Woodger, in *Logic, Semantics, Metamathematics*, 2nd edition, John Corcoran (ed.), 152–278. Indianapolis: Hackett, 1983.

Tarski, A., 1936, 'On The Concept of Logical Consequence', in Tarski 1983, 409—420.

Tarski, A., 1983, *Logic, Semantics, Metamathematics*, second edition, J. H. Woodger (trans.), John Corcoran (ed.), Indianapolis: Hackett.

Weir, A., 2003, 'Neo-Fregeanism: An Embarrassment of Riches', *Notre Dame Journal of Formal Logic*, 44/1: 13–48.

Whitehead, A.N., and B. Russell, 1910–1913, *Principia Mathematica*, 3 volumes, Volume 1 (1910), Volume 2 (1912), Volume 3 (1913), Cambridge: Cambridge University Press.

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*, Scots Philosophical Monographs, vol. 2, Aberdeen: Aberdeen University Press.

Zalta, E., 2006, "Essence and Modality", *Mind*, 115/459: 659–693.

Zalta, E., 2000, 'Neo-Logicism? An Ontological Reduction of Mathematics to Metaphysics', *Erkenntnis*, 53(1–2): 219–265.

Zalta, E., 1999, 'Natural Numbers and Natural Cardinals as Abstract Objects: A Partial Reconstruction of Frege's *Grundgesetze* in Object Theory', *Journal of Philosophical Logic*, 28(6): 619–660.

Zalta, E., 1988, *Intensional Logic and the Metaphysics of Intentionality*, Cambridge, MA: Bradford/MIT Press.

Zalta, E., 1983, *Abstract Objects: An Introduction to Axiomatic Metaphysics*, Dodrecht: D. Reidel.