A Defense of Logicism*

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

Uri Nodelman
Center for the Study of Language and Information
Stanford University
nodelman@stanford.edu

Edward N. Zalta
Center for the Study of Language and Information
Stanford University
zalta@stanford.edu

Abstract

We argue that logicism, the thesis that mathematics is reducible to logical and analytic truths alone, 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 a 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.

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 having 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. As part of our argument, we establish that the alternative notion of reduction that should be assumed by the thesis of logi-
A Defense of Logicism

cism gives up nothing important when it comes to establishing the most important goals set by the logicians 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 means. (Henceforth, we use ‘objects’ to refer generally to individuals, properties, and relations.) Thus, we agree with the early logicians 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, new logical objects will become definable. 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 \( \textit{bona fide} \) 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 logicians.

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 \( \textit{a priori} \) 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 on classic formulation:

But in reply to Kant, logicians claimed that these \( \textit{mathematical} \) propositions are \( \textit{a priori} \) 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 \( \textit{itself} \) 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’.  

\( \text{Benacerraf 1981, 42–43} \)

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 Godel’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 properties that they objectify or encode. 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 stateable in terms of primitive 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 (RM), etc. 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.

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 Null Set Axiom is expressed as $\neg \exists y (y \in \emptyset)$ instead of as $\exists x \neg \exists y (y \in x)$.

Now suppose that the language of ZF includes (higher-order) $\lambda$-expressions such as $\lambda x \varphi$, $\lambda F \varphi$, and $\lambda R \varphi$. In $\lambda x \varphi$, the $\lambda$ binds the individual variable ‘$x$’ to produce an expression that denotes a property of individuals; in $\lambda F \varphi$, the $\lambda$ 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 $\lambda$ binds the first-level 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

---

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.
are in argument position):\(^2\)

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

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

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

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) 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 \(\exists y (Sy \in \emptyset)\), 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) \(\text{ZF} \models \neg \exists y (Sy & y \in \emptyset)\)

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

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

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

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 ‘\(F\)’ as a variable ranging over second-level properties of first-level properties, and ‘\(R\)’ as a variable ranging over second-level properties of first-level relations, then:

\(^2\)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.

\((A)\) can be obtained by substituting the proposition \(\neg \exists y (Sy & y \in \emptyset)\) for \(p\) in ZF \(\models p\).

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

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

\((D)\) can be obtained by substituting \([\lambda R \neg \exists y (Sy & y R \emptyset)]\) for \(R\) in ZF \(\models R\in\).

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 ZF \(\models F\emptyset\); single out all and only those second-level properties \(F\) that satisfy the open formula ZF \(\models F\emptyset\) and \(\models F S\); and single out all and only those second-level properties \(R\) that satisfy the open formula ZF \(\models R\in\).

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 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 \(\varphi\). In particular, suppose that there is a unique abstract individual that can code up all of the first-level properties \(F\) such that 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 \(\iota x \varphi\), we could then formulate the following theoretical identification:

\[\emptyset_{2F} = \iota x (\iota! x & \forall F (xF \equiv F \emptyset_{2F}))\]

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 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 predicate \(F^n x_1 \ldots x_n\). 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, \(\emptyset_{ZF}\) is the hyperintensional object that encodes all and only the properties by which it is theoretically conceived, namely, all and only those \(F\)'s 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, \(ZF \models F_S\). 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 \(\varphi\). In particular, suppose that there is a unique first-level abstract property that can code up all of the second-level properties \(F\) that the property of being a set \((S)\) exemplifies in \(ZF\). Using ‘\(A!\)' now to denote the second-level property of being abstract, and \(F,F\) to assert that the first-level property \(F\) encodes the second-level property \(F\), we could then formulate the following theoretical identification:\(^4\)

\[
S_{ZF} = A!F & \forall F(F \equiv zF \models F_{ZF})
\]

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

\(^3\)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 \(\varphi\) (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 \(FF\), we’ve made the italic ‘\(F\)’ slightly smaller in size, so as to make it clear that \(F\) is the argument and \(F\) is the second-level property it encodes.

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

Finally, consider the last of the open formulas mentioned above, namely, \(ZF \models \mathcal{R}_e\). 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 \(\varphi\). In particular, suppose that there is a unique first-level abstract relation that can code up all of the second-level properties \(R\) that the membership relation \((\in)\) exemplifies in \(ZF\). Using ‘\(A!\)' to denote the second-level property of being abstract, and ‘\(\mathcal{R}\)' to assert that the first-level relation \(R\) encodes the second-level property \(R\), we could then formulate the following theoretical identification:\(^5\)

\[
\epsilon_{ZF} = A!R & \forall R(R \equiv zF \models R_{ZF})
\]

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

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 \(RR\), we’ve made the italic \(R\) slightly smaller in size, so as to make it clear that \(R\) is the argument and \(R\) 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:

\[ i \text{ is a type.} \]

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 \( \mathcal{L} \).

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.

**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 \( \Pi \) 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 \( n = 1 \), we read \( \Pi \tau_1 \ldots \tau_n \) as “\( \tau_1, \ldots, \tau_n \) exemplify \( \Pi \)”, and when \( n = 0 \), we read \( \Pi \) as “\( \Pi \) is true”. Truth is the 0-place case of exemplification.

**Encoding formulas.** There is one kind of encoding formula:

Where \( \tau \) is any term of type \( t \) and \( \Pi \) is a term of type \( \langle t \rangle \), then the expression \( \tau \Pi \) is an encoding formula.

We read \( \tau \Pi \) as: \( \tau \) encodes \( \Pi \).

**Complex formulas.**

Where \( \varphi, \psi \) are any formulas and \( \alpha \) is any variable, then \( \neg \varphi \) (‘it is not the case that \( \varphi \)’), \( \varphi \to \psi \) (‘if \( \varphi \), then \( \psi \)’) and \( \forall \alpha \varphi \) (‘every \( \alpha \) is such that \( \varphi \)’) are complex formulas.

We henceforth employ formulas of the form \( \varphi \land \psi, \varphi \lor \psi \), and \( \varphi \equiv \psi \), as these can be defined in terms of our complex formulas.

**Subformulas.** We define is a subformula of \( \varphi \) as follows:

1. \( \varphi \) is a subformula of \( \varphi \).
2. If \( \neg \psi \) is a subformula of \( \varphi \), then \( \psi \) is a subformula of \( \varphi \).
3. If \( \psi \to \chi \) is a subformula of \( \varphi \), then \( \psi \) and \( \chi \) are subformulas of \( \varphi \).
4. If \( \forall \alpha \psi \) is a subformula of \( \varphi \), then \( \psi \) is a subformula of \( \varphi \).
5. Nothing else is a subformula of \( \varphi \).

\footnote{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.}
We say that \( \psi \) is a proper subformula of \( \varphi \) just in case \( \psi \) is a subformula of \( \varphi \) but not identical to \( \varphi \).

**Propositional Formulas.** \( \varphi \) is a propositional formula iff \( \varphi \) 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 \( \alpha \) is any variable of type \( t \neq \langle \rangle \) and \( \varphi \) is any formula in which \( \alpha \) is the only variable that occurs free, then \( \lambda \alpha \varphi \) (“the \( \alpha \) such that \( \varphi \)”)) is a complex term having type \( t \).

2. **Complex relational terms.** Where \( \varphi \) 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 \( \lambda \alpha_1 \ldots \alpha_n \varphi \) (“\( \alpha_1, \ldots, \alpha_n \) such that \( \varphi \)”)) is a complex relation term having type \( \langle t_1, \ldots, t_n \rangle \), and (b) \( \varphi \) 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:

\footnote{This rules out two kinds of descriptions. First it rules out a description such as \( \psi(p \& \neg p) \). 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.}

\footnote{For purposes of this paper, we won’t need terms of the form \( \lambda \varphi \), where the \( \lambda \) 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 \( \epsilon \) is a constant of type \( \langle i, i \rangle \) and \( \mathcal{R} \) is a variable of type \( \langle \langle i, i \rangle \rangle \), then \( \mathcal{R}\epsilon \) is also an exemplification formula. If \( y \) is a variable of type \( i \), \( \emptyset \) is a constant of type \( i \) and \( \epsilon \) is a constant of type \( \langle i, i \rangle \), then \( y \in \emptyset \) is also an exemplification formula, given our convention of rewriting relatinional 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 \( xF, \mathcal{F}F \), and \( \mathcal{R}\mathcal{R} \). The formula \( xF \) is well-formed because \( x \) has type \( i \) and \( F \) has type \( \langle i \rangle \). The formula \( \mathcal{F}F \) 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 \( \mathcal{R}\mathcal{R} \), which again is well-formed because \( \mathcal{R} \) 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 \( \varphi \) if they are buried in a term within some propositional subformula of \( \varphi \). For example, the formula \( Rx(yG) (\langle x \rangle \land yG(yG)) \) exemplify the relation \( R' \) and the formula \( [\lambda xRx(xG)]z \) (‘\( z \) exemplifies the property of being an \( x \) that bears \( R \) to \( yG(yG) \)’) 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 \( \lambda \)-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 \( [\lambda x\exists F(xF \& \neg Fx)] \) are 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 \( \lambda \)-expressions such as \( [\lambda x p] \). These denote a property of individuals, i.e., a property with type \( \langle x \rangle \), 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 \( \lambda \)-expressions such as \( [\lambda y \varphi] \), \( [\lambda F \varphi] \), and \( [\lambda R \varphi] \), when \( \varphi \) 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 \( \lambda \) need not be free in \( \varphi \). 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 \( \varphi \), an individual \( x \) exemplifies \( [\lambda y \varphi] \) iff (the proposition denoted by) \( \varphi \) 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 \( o \) is a variable of type \( t \), we say that the property being ordinary (‘\( O! \)’) is being an \( o \) such that \( o \) fails to exemplify being abstract:

\[ O! =_{df} [\lambda o \neg Ao] \]

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:

\[
x = y =_{df} (O!x & O!y & \forall F (Fx \equiv Fy)) \lor (A!x & A!x & \forall F (Fx \equiv yF))
\]

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 \( H \) are all of type \( \langle \langle t \rangle \rangle \):

\[
F = G =_{df} (O! F & O! G & \forall x (Fx \equiv xG)) \lor (A! F & A! G & \forall H (F H \equiv G H))
\]

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 \( \langle (\rangle \rangle \).

I.e., where \( x \) is a variable of type \( i \) and \( O! \), \( A! \), and \( H \) have type \( \langle \rangle \):

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

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 \( \langle (t_1, \ldots, t_n) \rangle \). We leave the formal representation to a footnote.\(^9\)

\(^9\) Formally:

\[
F^{(t_1, \ldots, t_n)} = G^{(t_1, \ldots, t_n)} =_{df} (\text{where } n > 1)
\]

\[
O! F & O! G & \forall y^{t_2} \ldots \forall y^{t_n} ([\lambda x^{t_1} Fx^{t_1} y^{t_2} \ldots y^{t_n}] = [\lambda x^{t_1} Gx^{t_1} y^{t_2} \ldots y^{t_n}]) \lor \forall y^{t_1} y^{t_2} \ldots y^{t_n} ([\lambda x^{t_1} Fy^{t_1} x^{t_2} y^{t_2} \ldots y^{t_n}] = [\lambda x^{t_1} Gy^{t_1} x^{t_2} y^{t_2} \ldots y^{t_n}]) \lor \forall y^{t_1} \ldots y^{t_n} ([\lambda x^{t_1} F y^{t_1}, \ldots y^{t_n} x^{t_n}] = [\lambda x^{t_1} G y^{t_1}, \ldots y^{t_n} x^{t_n}]) \lor A! F & A! G & \forall H (F H \equiv G H)
\]

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:

\[
\text{Situation}(x) =_{df} \forall F (xF \to \exists p (F = [\lambda y p]))
\]

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 \( \Phi_{ZF} \) and \( \phi_{ZF} \) as terms of type \( i \), \( S_{ZF} \) as a term of type \( \langle i, i \rangle \), and \( \in_{ZF} \) of type of term \( \langle i, i \rangle \):

\[
\phi_{ZF} = \exists x (A! x & \forall F (xF \equiv zF \equiv F\phi_{ZF}))
\]

\[
S_{ZF} = \exists F (A! F & \forall F (FF \equiv zF \equiv F\phi_{ZF}))
\]

\[
\in_{ZF} = \exists R (A! R & \forall R (RR \equiv zF \equiv R\in_{ZF}))
\]

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 \( zF \equiv F\phi_{ZF}, zF \equiv F\phi_{ZF}, \) and \( zF \equiv R\in_{ZF} \), 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-

---

Hannes Leitgeb, Uri Nodelman, and Edward N. Zalta 18
modate the (negative) free logic of definite descriptions.\(^\text{10}\) Thus, a definite definition \(\alpha \varphi\) can be instantiated into universal claims only when it is known, by proof or by hypothesis, that \(\exists \beta (\beta = \alpha \varphi)\), i.e., that the description is logically proper.\(^\text{11}\)

2. The axiom governing the defined notion of identity.

3. The axioms governing the two kinds of complex terms: definite descriptions and the \(\lambda\)-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 \(t\) is not the power set of the domain of objects of type \(t\).\(^\text{12}\) 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 \(\alpha, \beta\) are distinct variables of the same type and \(\varphi'\) is the result of replacing zero or more free occurrences of \(\alpha\) in \(\varphi\) with occurrences of \(\beta\).

Thus, our defined 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 \(\beta\) is substitutable for \(\alpha\) in \(\varphi\).

This asserts: \(\beta\) is the \(\alpha\) such that \(\varphi\) if and only if \(\beta\) is uniquely \(\alpha\). As a simple example, let \(\alpha, \beta\) be the type \(i\) variables \(x, y\), respectively, let \(Q\) be a type \((i)\) constant, and let \(\varphi\) be the exemplification formula \(Qx\). Then the following is an instance of the Description Axiom:

\[
y = \iota x Qx \equiv \forall x (Qx = x = y)
\]

\(^11\)We shall assume familiarity with the following facts about free logic. First, the classical quantifier axiom for universal instantiation is modified so that terms \(\tau\) can only be instantiated into a universal claim if one knows that \(\exists \beta (\beta = \tau)\). Second, for every term \(\tau\) 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: \(\psi_{\iota \alpha \varphi} \rightarrow \exists \beta (\beta = \iota \alpha \varphi)\), where \(\psi\) is any atomic exemplification or encoding formula in which \(\alpha\) occurs as one of the arguments, \(\beta\) doesn’t occur free in \(\varphi\), and \(\psi_{\iota \alpha \varphi}\) is the result of substituting \(\iota \alpha \varphi\) for all the free occurrences of \(\alpha\) in \(\varphi\). This simply captures the principle underlying negative free logic that any atomic formula containing a non-denoting term is false.

\(^12\)In general, the domain \(D_t\) of type \(t\) is the union of the ordinary objects of type \(t\) and the abstract objects of type \(t\). So, the properties in \(D_{(t)}\) includes all the ordinary properties with type \((t)\) and the abstract properties with type \((t)\). It will be seen, upon inspection, that this is not the power set of \(D_t\).
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 \( \beta \)-Conversion for relations denoted by \( \lambda \)-expressions in which the \( \lambda \) binds one or more variables. To state \( \beta \)-Conversion, let \( \varphi^{\alpha_1,\ldots,\alpha_n} \) be the result of substituting the term \( \tau_i \) for every free occurrence of \( \alpha_i \) in \( \varphi \) \((1 \leq i \leq n)\):

\[
\text{AXIOM 2: } \left[ \lambda \alpha_1 \ldots \alpha_n \varphi \right] \beta_1 \ldots \beta_n \equiv \varphi^{\beta_1,\ldots,\beta_n}, \text{ provided } \beta_i \text{ is of the same type as, and substitutable for, } \alpha_i \text{ in } \varphi.
\]

This is just the familiar \( \beta \)-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:

\[
\begin{align*}
\text{(E) } & \forall G( [\lambda F \varphi] G \equiv \varphi_G^G ), \text{ provided } G \text{ is substitutable for } F \text{ in } \varphi \\
\text{(F) } & \forall S( [\lambda R \varphi] s \equiv \varphi_S^S ), \text{ provided } S \text{ is substitutable for } R \text{ in } \varphi
\end{align*}
\]

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 \( S_1 \):

\[
[\lambda F \neg \exists y(Fy \land y \in \emptyset)] S_1 \equiv \neg \exists y(S_1 y \land y \in \emptyset)
\]

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

\[
[\lambda R \neg \exists y(Fy \land y \in \emptyset)] \in \equiv \neg \exists y(Fy \land y \in \emptyset)
\]

Thus, instances of \( \beta \)-Conversion simply require that the denotation of the \( \lambda \)-expression to be a relation whose exemplification extension is equivalent to the \( \lambda \)-expression’s matrix. We’ll also assume that a principle of \( \alpha \)-Conversion, which asserts an identity between alphabetic variants, governs all our \( \lambda \)-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 \( \alpha \) 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:

\[
\text{AXIOM 3: } \exists \alpha(A! \alpha \land \forall F(\alpha F \equiv \varphi)), \text{ where } \varphi \text{ has no free } x_s.
\]

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 \( F \) are of type \( \langle i, i \rangle \). In the third, \( R \) is a variable of type \( \langle i, i \rangle \), while \( A! \) and \( R \) are of type \( \langle i, i \rangle \):

\[
\begin{align*}
& \exists x(A! x \land \forall F(xF \equiv \varphi)), \text{ where } \varphi \text{ has no free } x_F \\
& \exists F(A! F \land \forall F(FF \equiv \varphi)), \text{ where } \varphi \text{ has no free } x_F \\
& \exists R(A! R \land \forall R(RR \equiv \varphi)), \text{ where } \varphi \text{ has no free } x_R
\end{align*}
\]

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

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 $\varphi$, 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 $\varphi$ that have no free $x$s, $F$s or $R$s, respectively. In other words, the following are theorems:

$$\exists y (y = \iota x (A! x \& \forall F (xF \equiv \varphi)))$$
$$\exists G (G = \iota F (A! F \& \forall F (RF \equiv \varphi)))$$
$$\exists S (S = \iota R (A! R \& \forall R (R R \equiv \varphi)))$$

In what follows, we call such descriptions canonical since for any formula $\varphi$ (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 $\varphi$ and $\varphi'$ is the result of substituting $G$ for $F$ everywhere in $\varphi$.

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

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

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

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 $a^t_0$, is that abstract object of type $t$ that encodes all and only non-self-identical properties.

$$a^t_0 =_df \iota \alpha (A! \alpha \& \forall F (\alpha F \equiv F \neq F))$$

$a^t_0$ is indeed a null object since it provably encodes no properties. We may now assert, for any object $\alpha$ of type $t$, that if $\alpha$ is not identical to the the null object, then $\alpha$ is abstract if and only if $\alpha$ encodes a property. Where $\alpha$ is of type $t$, and $F$ and $A!$ are of type $\langle t \rangle$, this axiom may be formalized as follows:

**AXIOM 4**: $\alpha \neq a^t_0 \rightarrow (A! \alpha \equiv \exists \alpha F F)$

Note, as a simple corollary, that $\exists \alpha F F \rightarrow A! \alpha$, for if $\alpha$ 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 $\beta$-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 $\varphi$, where $\varphi$ is an exemplification pattern. But abstract properties and relations arise by comprehension, i.e., by what they encode, not by what exemplifies them. So $\lambda$-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:

**AXIOM 6:** $\neg A![\lambda \nu_1 \ldots \nu_n \varphi]$ \hspace{1cm} ($n \geq 1$)

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

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

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

### 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.

**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.  
  \[\text{[Said with reference to Zermelo-Fraenkel set theory.]}\]
- 2 is less than or equal to $\pi$.  
  \[\text{[Said with reference to real number theory.]}\]

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

---

\(^{14}\)Later, we will introduce special $\lambda$-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 $\beta$-Conversion doesn’t apply to them.

\(^{15}\)To give interested readers some hint, note that if we were to add modality to the present system, we could identify The Triangle (‘$\Phi_T$’) as the abstract individual that encodes exactly the properties necessarily implied by being triangular ($T$):

$$\Phi_T \overset{=_{df}}{=} \exists x (A!x \& \forall F (xF \equiv \square \forall y (Ty \rightarrow Fy)))$$

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 (‘$\approx_E$’):

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

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 $Gs$ (‘$\epsilon G$’) as the abstract individual that encodes exactly those $Fs$ that are materially equivalent to $G$:

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

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

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 $\lambda$-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 Zermelo-Fraenkel set theory, ‘$\mathcal{R}$’ denotes real number theory, ‘$\mathcal{PA}$’ denotes Peano arithmetic, etc. It follows that statements of the form $\mathcal{ZF} \models p$, $\mathcal{R} \models q$, $\mathcal{PA} \models r$, 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 $\varphi$ is a closed theorem of $T$, then $T \models \varphi^*$ shall be an axiom, where $\varphi^*$ 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:

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

the following statement will be an axiom of our framework:

$\mathcal{ZF} \models \neg \exists y (S_{\mathcal{ZF}} y \& y \in_{\mathcal{ZF}} \emptyset_{\mathcal{ZF}})$

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 $\lambda$-expressions, we also assume the principle of $\lambda$-Conversion, which has the form $[\lambda \alpha \varphi] \kappa \equiv \varphi^\kappa$, is part of the logic of $T$. Then our Importation Principle (by the logical closure of mathematical theories) yields the following special case, where $[\lambda \alpha \varphi^\kappa]$ is the result of substituting the variable $\alpha$ for all the occurrences of $\kappa$ in $\varphi$:

**Closure Principle.** Let $\varphi$ be a theorem of $T$, let $\kappa$ be a constant term of type $t$ occurring in $\varphi$, and let $\alpha$ be a variable of type $t$ that doesn’t occur free in $\varphi$, so that $[\lambda \alpha \varphi^\kappa]$ is a 1-place relation term of $T$ that has type (t). Then where $\varphi^*$ 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) $T \models ([\lambda \alpha \varphi^\kappa]^*_{T \mathcal{K} \mathcal{T}})_{T \mathcal{K} \mathcal{T}} \equiv \varphi^*$

(b) $T \models [\lambda \alpha \varphi^\kappa]^*_{T \mathcal{K} \mathcal{T}}$

i.e., (a) it is true in $T$ that: $\mathcal{K} \mathcal{T}$ exemplifies $[\lambda \alpha \varphi^\kappa]^*_{T \mathcal{K} \mathcal{T}}$ if and only if $\varphi^*$ is true in $T$, and (b) it is true in $T$ that $\mathcal{K} \mathcal{T}$ exemplifies $[\lambda \alpha \varphi^\kappa]^*_{T \mathcal{K} \mathcal{T}}$.

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

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

(G) $\mathcal{ZF} \models \neg \exists y (S_{\mathcal{ZF}} y \& y \in_{\mathcal{ZF}} \emptyset_{\mathcal{ZF}})$

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

(I) $\mathcal{ZF} \models [\lambda F \exists y (Fy \& y \in_{\mathcal{ZF}} \emptyset)]_{\mathcal{ZF}} S_{\mathcal{ZF}}$

(J) $\mathcal{ZF} \models [\lambda R \exists y (S_{\mathcal{ZF}} y \& y R_{\mathcal{ZF}})]_{\mathcal{ZF}} \in_{\mathcal{ZF}}$
Thus, not only does object theory become extended with axiom (G), i.e., in ZF, no set is a member of the null set, 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:

\[ ZF = \lambda x (A!x \land \forall F (xF \equiv \exists p(zF = p \land F = [\lambda y p])) ) \]

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

We can now say what the constants and \( \lambda \)-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 \( \lambda \)-expressions. Let \( \kappa_T \) be the \( T \)-indexed version of the primitive constant \( \kappa \) of \( T \), \( \Pi^1_T \) be the \( T \)-indexed version of the primitive 1-place predicate \( \Pi^1 \) of \( T \), \( \Pi^2_T \) be the \( T \)-indexed version of the 2-place predicate \( \Pi^2 \) of \( T \), and \( [\lambda \alpha^1 \varphi^2]_T \) be the \( T \)-indexed version of the 1-place \( \lambda \)-expression \( [\lambda \alpha^1 \varphi] \) 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:**

- \( \kappa_T = \lambda x (A!x \land \forall F (xF \equiv T \models F\kappa_T) ) \)
- \( \Pi^1_T = \lambda F (A!F \land \forall \bar{F} (\bar{F} \equiv T \models \bar{F}\Pi^1_T) ) \)
- \( \Pi^2_T = \lambda R (A!R \land \forall \bar{R} (\bar{R} \equiv T \models \bar{R}\Pi^2_T) ) \)
- \( [\lambda \alpha^1 \varphi^2]_T = \lambda \beta (A!\beta \land \forall \bar{\beta} (\bar{\beta} \equiv T \models \bar{\beta}[\lambda \alpha \varphi^*_T]), \text{ where } \beta \text{ is a variable of type } \langle t \rangle, A! \text{ has type } \langle \langle t \rangle \rangle, \text{ and } \Gamma \text{ is a variable of type } \langle \langle t \rangle \rangle. \)

The latter asserts that the \( T \)-based property being-an-\( \alpha^1 \)-such-that-\( \varphi^* \) is identical to the abstract property that encodes all and only those properties of properties \( \Gamma \) exemplified in \( T \) by the \( T \)-based property being-an-\( \alpha^1 \)-such-that-\( \varphi^* \). 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 \( \emptyset_{ZF} \), \( S_{ZF} \), and \( \in_{ZF} \). 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:**

- \( \emptyset_{ZF} = \lambda x (A!x \land \forall F (xF \equiv ZF \models F\emptyset_{ZF})) \)
- \( S_{ZF} = \lambda F (A!F \land \forall \bar{F} (\bar{F} \equiv ZF \models \bar{F}S_{ZF})) \)
- \( \in_{ZF} = \lambda R (A!R \land \forall \bar{R} (\bar{R} \equiv ZF \models \bar{R}\in_{ZF})) \)

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:

\[ [\lambda x \neg \exists y (S_{ZF} y y \in_{ZF} x)]_{ZF} = \lambda F (A!F \land \forall \bar{F} (\bar{F} \equiv ZF \models F[\lambda x \neg \exists y (S_{ZF} y y \in_{ZF} x)]_{ZF})) \]

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 \( \beta \) be a variable of type \( \langle \langle i \rangle \rangle \), let \( A! \) have type \( \langle \langle \langle i \rangle \rangle \rangle \), and let \( \Gamma \) be a variable of type \( \langle \langle \langle i \rangle \rangle \rangle \). Then the following is also a reduction axiom:

\[ [\lambda F \neg \exists y (F y y \in_{ZF} \emptyset_{ZF})]_{ZF} = \lambda \beta (A!\beta \land \forall \bar{\beta} (\bar{\beta} \equiv ZF \models \bar{\beta}[\lambda F \neg \exists y (F y y \in_{ZF} \emptyset)]_{ZF})) \]
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 $\beta$ be a variable of type $\langle \langle i, i \rangle \rangle$, let $\forall!$ have type $\langle \langle i, i \rangle \rangle$, and let $\Gamma$ be a variable of type $\langle \langle i, i \rangle \rangle$.

Then the following is also an instance of the Reduction Axiom schema:

$$\forall! (\forall X \neg \exists y (S_{x^2 y} \& y R_{x^2}) )_{\text{ZF}}$$

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:

**Equivalence Theorem Schemata:**

- $\forall G (\kappa_T G \equiv T \models G_{\kappa_T} )$
- $\forall G (\Pi^1_T G \equiv T \models G_{\Pi^1_T} )$
- $\forall S (\Pi^2_T S \equiv T \models S_{\Pi^2_T} )$

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

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

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 (I), (I), and (J) above may be instantiated, respectively, into the above claims to yield the following theorems:

- $\theta_{x^2} [\lambda x \neg \exists y (S_{x^2 y} \& y \in_{x^2} x)]_{x^2} \equiv \text{ZF} \models [\lambda x \neg \exists y (S_{x^2 y} \& y \in_{x^2} x)]_{x^2} \theta_{x^2}$
- $\phi_{x^2} [\lambda F \neg \exists y (F y \& y \in_{x^2} \theta_{x^2})]_{x^2} \equiv \text{ZF} \models [\lambda F \neg \exists y (F y \& y \in_{x^2} \theta_{x^2})]_{x^2} \phi_{x^2}$
- $\epsilon_{x^2} [\lambda R \neg \exists y (S_{x^2 y} \& y R_{x^2})]_{x^2} \equiv \text{ZF} \models [\lambda R \neg \exists y (S_{x^2 y} \& y R_{x^2})]_{x^2} \epsilon_{x^2}$

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 $\theta_{x^2}$, $\phi_{x^2}$, and $\epsilon_{x^2}$:

- $\theta_{x^2} [\lambda x \neg \exists y (S_{x^2 y} \& y \in_{x^2} x)]_{x^2}$
- $\phi_{x^2} [\lambda F \neg \exists y (F y \& y \in_{x^2} \theta_{x^2})]_{x^2}$
- $\epsilon_{x^2} [\lambda R \neg \exists y (S_{x^2 y} \& y R_{x^2})]_{x^2}$

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 theories in and of themselves, unprefix 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:

\[ \vdash_{\text{PA}} \text{N}0 \]

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

\[ \text{PA} \models \text{N}_{\text{PA}}0_{\text{PA}} \]

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, \( 0_{\text{PA}}N_{\text{PA}} \), is derivable as a theorem. We now have a precise and valid argument for why \( 0_{\text{PA}}N_{\text{PA}} \) is a true reading of “0 is a number”. No such argument can be given for the exemplification reading, a precise and valid argument for why \( \text{N}_{\text{PA}}0_{\text{PA}} \) 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 \( 0_{\text{PA}} \) encodes but also a true atomic fact about a property that \( N_{\text{PA}} \) encodes, namely, \( \lambda F 0_{\text{PA}}]N \). Once we import \( \vdash_{\text{PA}} \lambda F 0\]N so as to yield the axiom

\[ \text{PA} \models [\lambda F F0_{\text{PA}}]N_{\text{PA}}, \]

the Equivalence Theorem guarantees that the encoding formula \( N_{\text{PA}}[\lambda F F0_{\text{PA}}]_{\text{PA}} \) 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 \( N_{\text{PA}} \) makes \( \lambda F 0_{\text{PA}} \) is axiomatically false, by AXIOM 6 and the fact that \( \lambda F 0_{\text{PA}} \) 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):

\[ \text{ZF} \models \neg \exists y(S_{\text{ZF}}y \land y \in_{\text{ZF}} \emptyset_{\text{ZF}}) \]
\[ \text{ZF} \models [\lambda x \neg \exists y(S_{\text{ZF}}y \land y \in_{\text{ZF}} x)]_{\text{ZF}} \emptyset_{\text{ZF}} \]
\[ \text{ZF} \models [\lambda F \neg \exists y(Fy \land y \in_{\text{ZF}} \emptyset_{\text{ZF}})]_{\text{ZF}} S_{\text{ZF}} \]
\[ \text{ZF} \models [\lambda R \neg \exists y(S_{\text{ZF}}y \land y R \emptyset_{\text{ZF}})]_{\text{ZF}} \]

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:

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

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:

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

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, 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. Since we are focusing only on theoretical mathematics, logicism can be 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. 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:

---

\(^{17}\)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.

\(^{19}\)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 \(T\).
that each higher-order domain is not just non-empty but has at least two properties, the domain of individuals must be non-empty, the principle of

\( \lambda \)

ified the laws of classical logic from being considered logical. The existential import of
derive

claims from it, namely, the comprehension principle for relations. That is, one can derive

descriptions, and the principle of

\( \lambda \)

. But having existential import hasn’t disqualified the laws of classical logic from being considered logical. The existential import of

\( \lambda \)-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

\( \lambda \)-Conversion implies that each higher-order domain is not just non-empty but has at least two properties,

- 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

\( \lambda \)-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

\( \lambda \)-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 i) and being such that (represented by the \( \lambda \)).

Moreover, AXIOMS 4–7 (in Sections 3.3.1 and 3.3.2) stipulate what is meant by the property of

\( \text{being abstract} \)

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

\( \lambda \)-Conversion is a logical truth. Consider one of its instances, which is logically complex not only because it involves the \( \lambda \)-expression but also because it involves the negation symbol:

\[ [\lambda y \neg G y] x \equiv \neg G x \]

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

\( 20 \)

We recognize that

\( \lambda \)-Conversion has existential import; one can derive existence claims from it, namely, the comprehension principle for relations. That is, one can derive \( \exists F \forall x (F x \equiv \varphi) \) 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

\( \lambda \)-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

\( \lambda \)-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

\( \lambda \)-Conversion make minor existence demands.
form “x is a non-G exemplifier” ([λy ¬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” ([λy 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 [λy Gy & Hy]x ⊢ Gyx). The point also applies to relations and relational λ-expressions. Complex reasoning about the converse of relation R ([λxy Ryx])21 and relations like unrequited love ([λxy Lxy & ¬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

Under this conception, not only is λ-Conversion a logical truth, but the inference from ¬Gx to [λy ¬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

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.
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 $\lambda$-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 $\lambda$-Conversion and higher-level abstracta. For take the example:

$$\lambda R \neg \forall x Rx x S \equiv \neg \forall x S x x$$

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 $\lambda R \neg \forall x Rx x$, 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 higher-order properties of relations, but instead of representing this pattern as a property of type $\langle \langle i, i \rangle \rangle$ (property of properties of relations), encoding predication turns that pattern into an abstract relation of type $(i, i)$ that encodes the properties of relations $R$ that satisfy the pattern $T_D \models R <_D$. (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 $\varphi$ containing the term $<_D$ is logically true if and only if $\varphi$ is true in all interpretations that include all the (abstract) objects required for the possibility of having the thought that $\varphi$. So $<_D$ must be contained in the domain of interpretation. Thus, the following existence claim is logically true:

$$\exists R (A ! R \& \forall R (R R \equiv T_D \models R <_D))$$

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:

$$\exists R (A ! R \& \forall R (R R \equiv T \models R S))$$
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$_{PA}$, of the abstract relation property successor$_{PA}$, and the abstract individual Zero$_{PA}$ 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

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.

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.

26This 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 $\varphi$ and $\varphi \rightarrow \psi$ are logical in virtue of being required for the possibility of abstract thought and reasoning, then $\psi$ 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 “θ ∈ {θ}” is analytic, and yet still claim that “In ZF, θZF ∈ ZF {θ}ZF” is. The latter is true in virtue of the meaning of the terms ‘ZF’, ‘θZF’, ‘∈ZF’, and ‘{θ}ZF’. Since ‘ZF’ denotes a theory, and a theory encodes its theorems, ‘In ZF, θZF ∈ ZF {θ}ZF’, when represented as ZF |= θZF ∈ ZF {θ}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 of a mathematical term is identical to the meaning of the canonical description used to identify the denotation of that term. 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:

\[
\begin{align*}
\thetaZF & = tR(A!x & \forall F(xF \equiv zF \models F\thetaZF)) \\
\sigmaZF & = tF(A!F & \forall F(FF \equiv zF \models F\sigmaZF)) \\
\epsilonZF & = tR(A!R & \forall R(R\epsilon \equiv zF \models R\epsilonZF))
\end{align*}
\]

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 \(c_\tau\).

(P2) The mathematical concept \(c_\tau\) 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 \(c_\tau\) 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 \(\tau'\) are identical, then τ = \(\tau'\) 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 mathematic 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 \(c_\tau\) 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, \(c_\tau\) can be identified with the inferential role of τ in T. For example, \(c_\theta\), \(c_\pi\), \(c_\epsilon\), \(c_c\), etc., are identical to the inferential roles of \(\theta\), \(\pi\), \(\epsilon\), \(<\), 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. 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 ‘∅’ in ZF is properly identified by the canonical description, \( \land_x (\forall x \& \forall F (xF \equiv \exists F \in ZF)) \). This holds because the abstract object denoted encodes all and only the properties of the null set derivable in ZF. For instance, it is derivable in ZF that \( \emptyset \in \{\emptyset\} \). This theorem gets imported into object theory as \( \forall \emptyset \in \{\emptyset\} \). So by the Closure Principle, it follows that \( \forall \emptyset \in \{\emptyset\} \emptyset \). As such, the property \( \{\emptyset \in \{\emptyset\}\} \) is one of the properties that satisfies the formula \( \forall \emptyset \in \{\emptyset\} \emptyset \). 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 meanings 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.

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

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.

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

---

27See, 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.

28Again, 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.
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:

\[ \text{LC'} \quad \text{The meaning of a mathematical term } \tau, \text{ i.e., the mathematical concept } c_\tau, \text{ is identical to a logical object.} \]

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 \( \tau \) 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 \( \text{ZF} \) to \( \text{ZFC} \), 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 \( \text{ZF} \) and \( \text{ZFC} \). Note that the Platonist claim presupposes both that sets exist independently of our theories of them and that when we move from \( \text{ZF} \) to \( \text{ZFC} \), the new theory is simply characterizing the objects of \( \text{ZF} \) 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 ‘\( \mathcal{E} \)’, for that body of truths and replace ‘\( \text{ZF} \)’ in our reduction axioms from above by ‘\( \mathcal{E} \)’. Then everything should go as before, and we should be able to reconstruct the mathematical terms of set theory as logical expressions. That is, we can plug \( \mathcal{E} \) 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 \( \text{ZF} \) and \( \text{ZFC} \) will be true of the denotations that we have assigned to the terms and predicates of set theory \( \mathcal{E} \) since \( \mathcal{E} \) includes these theorems (assuming Choice is included in \( \mathcal{E} \)). 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

---

\(^{32}\)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.

\(^{33}\)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 ‘embarrassment 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 company objection, the embarrassment 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 commitments, 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 \(\exists x \, x = y\)). 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 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 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.
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

[In preparation.]

Bibliography


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


