SECOND-ORDER LANGUAGES AND MATHEMATICAL PRACTICE

STEWART SHAPIRO

There are well-known theorems in mathematical logic that indicate rather profound differences between the logic of first-order languages and the logic of second-order languages. In the first-order case, for example, there is Gödel's completeness theorem: every consistent set of sentences (vis-à-vis a standard axiomatization) has a model. As a corollary, first-order logic is compact: if a set of formulas is not satisfiable, then it has a finite subset which also is not satisfiable. The downward Löwenheim-Skolem theorem is that every set of satisfiable first-order sentences has a model whose cardinality is at most countable (or the cardinality of the set of sentences, whichever is greater), and the upward Löwenheim-Skolem theorem is that if a set of first-order sentences has, for each natural number \( n \), a model whose cardinality is at least \( n \), then it has, for each infinite cardinal \( \kappa \) (greater than or equal to the cardinality of the set of sentences), a model of cardinality \( \kappa \). It follows, of course, that no set of first-order sentences that has an infinite model can be categorical. Second-order logic, on the other hand, is inherently incomplete in the sense that no recursive, sound axiomatization of it is complete. It is not compact, and there are many well-known categorical sets of second-order sentences (with infinite models). Thus, there are no straightforward analogues to the Löwenheim-Skolem theorems for second-order languages and logic.¹

There has been some controversy in recent years as to whether “second-order logic” should be considered a part of logic,² but this boundary issue does not concern me directly, at least not here. The present approach is to assess the adequacy of first-order languages in formalizing actual mathematical practice. This problem is one that occupied mathematicians and logicians earlier this century (see Moore [1980]), but seems to have received less attention recently. My main conclusion, in agreement with Bernays, Hilbert, and Zermelo (and in disagreement with Gödel and Skolem), is that no first-order language is sufficient for axiomatizing such branches

¹ A detailed exposition of the technical differences between first-order logic and second-order logic is found in Boolos and Jeffrey [1980, Chapter 18].
² Quine [1970] suggests that the set-theoretic presuppositions of (the standard semantics of) second-order logic actually disqualify it as part of logic. On the other hand, Corcoran [1973] and, more extensively, Boolos [1975] argue against this, that second-order logic is logic. From a different perspective, Tharp [1975] suggests that since first-order logic has certain properties (such as completeness and compactness) one would want a logic to have, ceteris paribus it is preferable.

© 1985, Association for Symbolic Logic
0022-4812/85/5003-0015/$03.90

714
as arithmetic, real and complex analysis, and set theory—branches that (each) deal
with a particular (infinite) structure or, in other words, branches whose languages
have "intended interpretations". The argument presented rules out any language
whose logic is either complete or compact. I go on to suggest that nothing short of a
language with second-order variables will do.

The considerations brought against first-order languages are semantical. That is
to say, I argue that the semantics of first-order languages is not adequate for the pre-
formal semantics of mathematical practice. A few brief comments elaborating this
perspective are in order.

First, the standpoint of this article might best be characterized as a "neutral
realism". The "realism" indicates that mathematical discourse is taken at face value.
Contra formalism, (most) mathematical assertions are regarded as meaningful
assertions about mathematical entities. Mathematical truth is determined by the
subject matter of mathematics and, thus, "truth" is synonymous with neither
(real/ideal) "knowledge" nor "provability". Contra intuitionism and logicism, no
attempt is made to criticize the bulk of mathematical practice. Rather, actual
mathematical practice is taken to be the data for the considerations of this paper.
The "neutral" in "neutral realism" indicates that at present, I have no view on the
makeup or the ontological status of the subject matter of mathematics. I only hold
that there is such a subject matter. In fact, I have no a priori objection to any
interpretation of mathematics as long as the integrity of the bulk of mathematical
discourse is preserved.3

It might be noted that much (but not all) of the literature on both sides of the first-
order/second-order issue takes or presupposes a viewpoint of realism. For example,
Gödel, an avowed platonist, was one of the strongest (and probably the most
influential) proponents of first-order languages. Moreover, the recent arguments
(see footnote 2) against higher-order languages are directed at the semantics of such
languages and, consequently, seem to presuppose a realism of sorts. Indeed, if one is
concerned only with codifying mathematical proof (for example, if one is a formalist
of the Hilbert or Curry school), then there is little to object to concerning any
effectively specified language and deductive system. In short, if one is not concerned
with the interpretation of a language of mathematics as such, then, a fortiori, one will
not worry about such things as excess ontological commitment and inconvenient
semantic properties.

Higher-order languages are considered here with standard semantics in which, for
a given interpretation, the second-order predicate variables range over all of the
subsets of the domain, the second-order function variables range over all of the
functions from the domain to the domain, etc. There is, of course, an alternate
semantics for second-order languages, developed originally in Henkin [1950], in
which, for a given interpretation, the predicate variables range over a fixed subset
of the power-set of the domain, etc. For this alternate semantics (and the usual
deductive system), second-order logic is sound, complete, and compact. Although it
will not always be demonstrated directly, it is easily seen that most of the present

3 Resnik [1980, Chapter 5] uses the term “methodological platonist” to refer to a similar, if not
identical, position.
considerations against first-order languages apply to second-order languages with Henkin semantics.

Finally, the completeness theorem indicates that for first-order languages, the proof theory corresponds in a direct way with the semantics, or model theory. Consistency and satisfiability are coextensive, as are deductive consequence and semantic consequence. This, of course, is not the case with second-order languages. Thus, present considerations concerning the semantics of mathematics do not shed much light on the question of which deductive systems are appropriate for codifying mathematical proof. To put it differently, my thesis is that reference to the predicates or subsets of given domains is necessary to capture the semantics of mathematical practice, but I have little to say concerning the particular axioms or assumptions about such subsets necessary to codify normal proof techniques.4

As noted, first-order languages do not allow categorical characterizations of infinite structures. I take this as their main shortcoming. §1 deals with the importance of categoricity in understanding and communicating mathematics. This involves the relevance of the Lüwenheim-Skolem theorems to the present issue and the epistemic presuppositions of second-order languages. §2 concerns further inadequacies of first-order versions of arithmetic, analysis, and set theory, concluding that such theories do not capture important, perhaps crucial, aspects of those fields. §3 is a discussion of the adequacy of several alternate languages. The first subsection concerns languages “intermediate” between first-order and second-order; the second subsection concerns the language of first-order set theory. The final §4 is a brief comparison of the semantics and proof-theoretic strength of standard first-order logic with that of standard second-order logic.

One of the purposes of logic is to codify correct inference. Thus, if my major conclusions are correct, the underlying logic of many branches of mathematics is (at least) second-order: one cannot codify the correct inferences of a second-order language with a first-order logic. It follows that the inconvenient technical properties and presuppositions of second-order logic must be accepted. The correct conclusion, I believe, is that there is no sharp distinction between logic and mathematics. The study of correct inference, like almost any other science, involves some mathematics and some mathematical presuppositions.

§1. Categoricity and the Lüwenheim-Skolem theorems. In broad terms, one major purpose of axiomatizing a branch of mathematics is to codify the practice of that branch. Historically, this has two distinct, but related aspects, one involving the deductive system of the language of axiomatization, the other the semantics.

One purpose of axiomatization is to organize and systematically present the truths and correct inferences of the branch.5 The goal of this aspect is completeness:

4 The problem concerning the proof theory of mathematical practice is treated in some detail in Feferman [1977].

5 As is well known, the organization of mathematical assertions was one of the chief aims of the Hilbert program:

When we are investigating the foundations of a science, we must set up a system of axioms which contains an exact and complete description of the ... ideas of that science ... no statement within the realm of the science ... is held to be true unless it can be deduced from the axioms by means of a finite number of logical steps. (Hilbert [1900, Problem 2])
an axiomatization—language and deductive system—is complete iff it has as theorems all (and only) the truths of that branch. Of course, the incompleteness theorem (and subsequent work in the theory of computability) shows this goal to be unattainable even for arithmetic in any suitable language.

The other purpose of axiomatization is to describe a particular structure, an intended interpretation of a branch of mathematics. At least one goal of this aspect is categoricity: an axiomatization—language and semantics—is categorical iff any two of its models are isomorphic. One of the first writers to discuss categoricity was Oswald Veblen in his axiomatization of geometry:

In as much as point and order are undefined, one has a right, in thinking of the propositions, to apply the terms in connection with any class of objects of which the axioms are valid propositions. It is part of our purpose, however, to show that there is essentially only one class of which the ... axioms are valid ... a system of axioms is categorical if it is sufficient for the complete determination of a class of objects or elements. (Veblen [1904, 346–347])

In this passage, Veblen seems to take categoricity as the primary aim of his axiomatization. This is underscored with a remark that if one has a categorical axiomatization, then “any further axiom would have to be considered redundant”, to which a footnote is added “... even were it not deducible from the [other] axioms by a finite number of syllogisms” (p. 346). A potential axiom (i.e., a true sentence) which is not deducible from the other axioms can be considered redundant only for the purpose of describing a structure. Such a sentence is certainly not redundant for the purpose of organizing and presenting the truths of a branch of mathematics.

The role of categoricity in the history of mathematics and logic is discussed in Corcoran [1980]. For present purposes, a simple, but important, point is that this aspect of the enterprise of axiomatization involves a distinction between a mathematical structure itself, and the language used to describe a structure. It is clear (at least with hind-sight) that if an axiomatization correctly describes a structure, then it also correctly describes any isomorphic structure. Thus, for the purpose of description, a categorical axiomatization is the best one can do. In Corcoran’s words:

The insight that truth in a formal language depends solely on the form of the interpretation (and is independent of content ...) is partly reflected in

---

6 A similar distinction applies to the enterprise of logic (and metalogic) and roughly corresponds to the distinction between proof theory and model theory. Let $L$ be a language. One purpose of the logic of $L$ is to codify the correct inferences of $L$. Gödel’s completeness theorem shows that this purpose can be achieved for first-order languages and, as noted, this purpose cannot be completely achieved for second-order languages (with standard set-theoretic semantics in both cases). A second purpose of the logic of $L$ is to describe what correct inference in $L$ amounts to. For the languages presently under consideration, this involves an account of the set-theoretic semantics. I suggest that the second purpose does not depend on the first—if anything, it is the other way around. That is, I maintain that (for second-order languages in particular) it is possible to characterize “correct inference” without being able to determine the correctness of every proposed inference. This stance concerning logic is part of my “neutral realism”, but it seems to be shared by several nominalists. (See Field [1980] and Gottlieb [1980].)
the fact that isomorphic interpretations have the same same set of truths ...
Moreover, it has been clear since the turn of the century ... that given any interpretation i, there are other interpretations isomorphic with i but having no content in common with i. The existence of such isomorphic "images" implies, of course, the impossibility of uniquely characterizing an interpretation by means of a set of sentences in a formal language. Accordingly, it is sometimes said that the best possible characterization of an interpretation would be a "characterization up to isomorphism"... (Corcoran [1980, 190])

Of course, the Löwenheim-Skolem theorems imply that no set of sentences in a first-order language can be a categorical description of an infinite structure. Thus, for example, any first-order axiomatization A of the natural numbers has "unintended" interpretations in the sense that there are models of A which are not isomorphic to the natural numbers. Thus, with a first-order language, a completely successful description of the natural numbers is not possible. Moreover, it is not difficult to see that a language and logic that does contain a categorical axiomatization of arithmetic is neither compact nor complete. Similar remarks apply to any infinite structure.

As is well known, much has been written on the philosophical significance of the Löwenheim-Skolem theorems. A general survey of this literature would go well beyond the scope of the present paper, but the remarks of two authors, Skolem [1923] himself and John Myhill [1951], are of particular interest.

Skolem's [1923] celebrated conclusion is that all set theoretic "notions" are "unavoidably relative". For example, one cannot claim that a given domain D is uncountable simpliciter, but only uncountable relative to a certain model (containing D) of set theory. For any such D, one cannot rule out the possibility that D may be countable relative to a (richer) model, one that contains a function from the natural numbers onto D. For a second example, Skolem holds that this relativity applies to the Dedekind notions of "finite" and "simply infinite sequence" (or model of the natural numbers). It seems that if one accepts the modern trend of regarding all mathematical notions as set-theoretic, then Skolem's conclusion entails the relativity of virtually all mathematical notions, including those of natural number and real number. Thus, it appears that Skolem rejects the above distinction between a mathematical structure and a formal axiomatization and, in particular, rejects the notion of an intended interpretation of an axiomatization.

The conclusions of Myhill [1951] are less extreme. He accepts the distinction between structures themselves and formal languages involved in description, but he argues that the Löwenheim-Skolem theorems indicate an inadequacy in the enterprise of formalization. His argument begins with the suggestion that mathematical structures are first apprehended by "intuition". Part of the purpose of formal axiomatization is to describe, communicate, and thus help study a particular structure. Myhill agrees that although deducing the properties of a structure only requires a sound axiomatization, the description and communication of a structure is possible only to the extent that an axiomatization uniquely characterizes it. He claims, however, that the Löwenheim-Skolem theorems show that for infinite structures, this is impossible. His first conclusion is that mathematicians cannot
dispense entirely with the original intuitions which determined the intended interpretation of the axiomatization:

[In a formalism] we operate with symbols which keep their shape rather than with ideas which fly away from us. All real mathematics is made with ideas, but formalism is always ready in case we grow afraid of the shifting [of our ideas]. The Skolem "paradox"... proclaims our need never to forget completely our intuitions. (Myhill [1951])

If this analysis is correct, then our ability to use a formalism to understand others is somewhat limited. Even if two mathematicians agree on an axiomatization of, say, arithmetic, real analysis, or set theory, they cannot be sure that they have in mind the same (or even isomorphic) interpretations of their agreed-on axiomatization. Concerning set theory, Myhill wrote:

... there seems to be no formal means of assuring that our concept of membership is the same as another person's... For no... number of formal assertions agreed on by us both could be evidence that his set-theory was not in my sense denumberable... The second philosophical lesson of the Löwenheim-Skolem theorem is that the formal communication of mathematics presupposes an informal community of understanding.

Both Skolem and Myhill thus speak of the limits of formalization and formal languages. Since the Löwenheim-Skolem theorems only apply to first-order axiomatizations, their considerations seem to presuppose that all legitimate formalizations of mathematical practice employ first-order languages. I propose here that their arguments actually represent a reductio ad absurdum against this presupposition.

There is virtually universal agreement among mathematicians that arithmetic, real analysis, and complex analysis each deals with a single, specific structure—the intended interpretation of the axiomatization. There is also some (but not universal) agreement that the same holds for set theory. Of course, the philosophical issues at hand are not to be settled by popularity, but this data is striking. Those working in arithmetic, for example, know that any first-order axiomatization has nonstandard interpretations, yet they believe both that there is a standard interpretation of arithmetic (or, at any rate, a class of standard interpretations, all of which are isomorphic) and that every other mathematician has in mind the same (or an isomorphic) interpretation. I take it as undisputed that every mathematician does have in mind the same (or an isomorphic) structure of natural numbers. Similar remarks apply to real analysis and complex analysis, if not set theory.

A question naturally arises as to how these structures are apprehended and communicated. Several philosophers (such as Myhill, as above, and Gödel) suggest that at least some mathematical structures are apprehended through a faculty of intuition. Postulating or suggesting such a faculty, however, does not completely solve the present problems. At best, a faculty of intuition can account for how a single mathematician apprehends, say, the natural number structure and then

---

7 See Moore [1980] for an extended discussion of Skolem's views on this issue.
describes it in a formal or informal language. A question remains as to how a second mathematician could know which structure the first is describing. Presumably, she also has a faculty of intuition, but she can just as well intuit several structures including, perhaps, that of a nonstandard model of arithmetic. The problem is one of communication.

Myhill’s point is that (first-order) formal languages are inadequate to insure communication. He suggests that the fact that mathematical structures are communicated presupposes an “informal community of understanding”. This underscores the present problem but does not solve it.

One could, I suppose, postulate a faculty of mental telepathy between mathematicians to account for the communication of structures; but, without this, all communication is mediated by language. This is where categoricity is important.

To reiterate, different mathematicians do understand “the natural number structure”, “the real number structure”, and “the complex number structure” the same way (or, at least, in isomorphic ways). That is, mathematicians succeed in communicating these structures to each other. The informal language of mathematics is thus sufficient to insure this communication. Second, the purpose of formal axiomatization is to codify mathematical practice, one of whose purposes is the description and communication of structures. I conclude that a language and semantics of formalization should be sufficient to insure this communication. That is, the language of formalization should allow categorical characterizations. It follows that first-order axiomatizations are inadequate.

A related conclusion is found in Montague [1965]; he shows that the notion of a “standard model” of arithmetic, real analysis, or even set theory can be understood as a model of the respective second-order theory. Thus, it is proposed that second-order languages are sufficient to insure the description and communication of these structures.8

This proposal, however, can be challenged and, in response, must be qualified. The assertion that a given structure is described (up to isomorphism) and communicated by a second-order language depends on a premise that the second-order language is itself unambiguously understood. That is, the categoricity of the second-order theories in question depends on there being a unique and clearly understood interpretation of such second-order quantifiers as “all subsets”. Let T be a second-order theory and D a structure with domain d. The statement that D is a model of T has quantified variables ranging over the (entire) collection of subsets of d. It is thus possible (perhaps) for two mathematicians to disagree whether D is a model of T if they have in mind different “powersets” of d.9 It is thus conceded that

---

8 §3 below concerns the extent to which second-order languages are also necessary for the description and communication of structures.

9 Similar remarks apply to the collections of relations on d and functions from d to d. This, in effect, is the insight behind the Henkin semantics for second-order languages. This objection was suggested to me by Mike Resnik and Nicolas Goodman. A similar consideration is found in Weston [1976].

Actually, the only role of the second-order terminology in standard claims and proofs of categoricity is to apply prenex universally quantified sentences to particular subsets (which may not be definable in the basic first-order theory alone). Thus, to grasp, understand, and follow the claims of categoricity, one need only grasp this kind of inference. More on this below.
the proposal concerning the expressive ability of second-order languages presupposes that the second-order quantifiers in the relevant definitions are understood unambiguously. A few remarks are in order.

First, understanding the second-order quantifiers of a given theory is not the same as grasping the set-theoretic hierarchy. In a given theory, the quantifier "all subsets" ranges over the collection of subsets of a fixed domain. In general there is no powerset operator to be iterated. Second-order arithmetic, for example, only has variables ranging over natural numbers and sets (and relations) of natural numbers. There are no variables ranging over such items as sets of sets of numbers and functions from numbers to sets of numbers. The set-theoretic hierarchy, on the other hand, is a proper class that contains the result of iterating the powerset operator into the transfinite.

I would suggest, then, a distinction between the "logical" conception of set and the "iterative" conception of set. The iterative conception refers to the set-theoretic hierarchy, itself an extremely large mathematical structure. The logical conception, on the other hand, occurs only in the context of a fixed domain. Thus, a second-order theory of a given domain has variables ranging over the collection of logical sets vis-a-vis that domain. With this terminology, the presupposition at hand is that for a fixed domain, the second-order quantifier "all logical subsets" is unambiguously understood.¹⁰

As above, let $T$ be a theory and $D$ a structure with domain $d$. I submit that what it means for a collection $c$ to be a (logical) subset of $d$ is clear and unambiguous: $c$ is a subset of $d$ if and only if every member of $c$ is a member of $d$. What is at issue here is whether the totality (or range) of subsets of $d$ is itself clear and unambiguous. Let $P_1$ and $P_2$ be two candidates for the range of the second-order predicate quantifiers of $T$ (via-à-vis $D$). That is, let $P_1$ and $P_2$ be two candidates for the logical powerset of $d$. I suggest that if $P_1 \neq P_2$, then there is a clear sense in which (at least) one of them is not the powerset of $d$. Indeed, suppose that there were a collection $c$ such that $c \in P_1$ but $c \notin P_2$. I take it that (for a classical mathematician) it is determinate whether

¹⁰ There is a tradition, originating, perhaps, with Boole and including Peirce and Schroder, which takes the subsets of a domain to be under the purview of logic. Moreover, much of Gödel [1944] is a defense of the claim that mathematical logic deals with "classes, relations, . . . instead of numbers, functions, geometrical figures, etc." (reprint, p. 211). This tradition, of course, predates the development of the set-theoretic hierarchy (as a subject of mathematics itself) by Cantor, Zermelo, etc. (See Moore [1980, especially 97–101].) Cantor's distinction between sets and "inconsistent totalities" may also be closely related to the present distinction between iterative sets and logical sets (in the context of set theory). See Wang [1974, Chapter VI] for a detailed discussion of the historical distinction between the "limitation of size" doctrine of sets and the logical doctrine of sets. One may also compare a distinction made by Frege and (the early) Russell between sets-as-composed-of-their-elements and classes-as-extensions-of-concepts. (The differences between these authors notwithstanding. See, for example, Chapter VI and Appendix A of Russell's Principles of Mathematics.)

In the sequel, when the word "set" is employed, the context usually indicates whether "logical set" or "iterative set" is meant. An ambiguity can result only in the context of set theory because, in this case, the fixed domain in question (vis-à-vis the logical conception) is the set-theoretic hierarchy. In other words, in set theory, a "logical set" is a collection of iterative sets. Here I follow the custom of designating such collections classes. The upshot of Russell's paradox is that in the context of set theory, there are logical sets that are not iterative sets.
every element of $c$ is an element of $d$. If every element of $c$ is in $d$, then $P2$ is not the powerset of $d$; otherwise, $P1$ is not the powerset of $d$.

Finally, my presupposition/suggestion that the locution “all logical subsets” is unambiguously understood entails only that when a mathematician uses the phrase in connection with a domain, he and his listeners understand what he means the same way. I do not make the absurd claim that any or all of the properties of the powerset (such as its cardinality, or whether it contains a nonconstructible element) are known.

§2. First-order axiomatizations. This section deals with specific inadequacies of first-order axiomatizations of branches of mathematics. §2.1 concerns three concepts—finitude, minimal closure, and well-foundedness—which form an important part of general mathematical practice, but which cannot be formulated in first-order languages. §2.2 is a comparison of the standard first-order versions of arithmetic, real analysis, and set theory with their second-order counterparts.

2.1. Let $T$ be a first-order theory and $\Phi$ a formula with one free variable. For each natural number $n$, there is, of course, a first-order formula that asserts that the extension of $\Phi$ (or the domain of discourse of $T$) has at most $n$ members. There are, however, circumstances in which one wishes to assert that a given extension (or domain) is finite without specifying a fixed (numerical) bound on the cardinality of this extension. Consider, for example, the theory of finite groups in which it is stated that the domain of discourse is finite, or the theory of computability in which each Turing machine is characterized by a finite number of instructions (but it being crucial that for each natural number $n$, there are Turing machines with at least $n$ instructions).

I submit that finitude is a clear and unambiguous concept. If, for example, a mathematician asserts that a given extension is finite, then his listeners understand what he means. Therefore, a language used to formalize mathematical practice must be capable of expressing this property. No first-order language can do this. As is well known, if, for each natural number $n$, there is a model of $T$ in which the extension of $\Phi$ (or the domain of discourse) has at least $n$ members, then there is a model of $T$ in which the extension of $\Phi$ (or the domain of discourse) is infinite.

Notice that the second-order formula

$$\forall f [(\forall x(\Phi(x) \rightarrow \Phi(fx)) \& \forall y \forall z(fy = fz \rightarrow y = z)) \rightarrow \forall y(\Phi(y) \rightarrow \exists x(fx = y))]$$

is satisfied by all those, and only those, models of $T$ in which the extension of $\Phi$ is finite.

In a similar vein, Boolos [1981] shows that theories formulated in first-order languages cannot express such simple cardinality comparisons as “the extension of $\Phi$ is at least as large as the extension of $\Psi$”.

Moving to the second example, it is common to describe a particular set (or structure) through what may be called a minimal closure property. The construction is usually carried out in a background theory $T$ concerning a model $M$ with domain $d$. To describe a set $A \in d$ by this technique, one first gives a basis subset $B \in d$ and a set $F$ of functions or operations (or, perhaps, relations) on $d$. The set $A$ is then characterized as the “smallest” set which contains $B$ and is closed under the
operations in $F$. The set $A$ is sometimes described as the set obtained from $B$ by closure under the operations in $F$, or one says that $y$ is in $A$ just in case $y$ is the result of an iteration of the operations in $F$ on members of $B$ finitely many times.

There are numerous examples of the minimal closure construction. To mention a few, if $R$ and $S$ are two rings, $R \subseteq S$ and $a \in S$, then the ring $R[a]$ can be described as a minimal closure (within $S$) with basis $R \cup \{a\}$ and the functions of $S$-addition and $S$-multiplication. The rational subfield of a field is a minimal closure with basis $\{1\}$ under the field functions and their inverses. In analysis, the statement that the real numbers are Archimedean amounts to a claim that for every positive real number $r$, the minimal closure of $\{r\}$ under addition is unbounded. In proof theory, collections of terms, well-formed formulas, and theorems are defined as minimal closures on sets of strings, and, in model theory, elementary submodels are often constructed as minimal closures under a set of Skolem functions.

From the practice of mathematicians, I submit that the use of minimal closure is well-understood and that there is no ambiguity concerning the constructed set. It is, therefore, a requirement on languages of axiomatization that they be capable of expressing minimal closures. As with finitude, a straightforward compactness argument shows that no collection of first-order formulas can successfully define any nontrivial minimal closure.

Notice that the following second-order formula $\theta(x)$ does characterize the minimal closure of the extension of $P$ under the function denoted by $f$:

$$\forall X \{\forall y[(\Phi(y) \rightarrow Xy) \& (Xy \rightarrow Xfy)] \rightarrow Xx\}.$$  

That is, $\theta(x)$ holds (in any model of the background theory) just in case $x$ is in the minimal closure of the extension of $\Phi$ under the function denoted by $f$.

The final example is that of a well-founded relation. A binary relation $E$ is well-founded iff there is no infinite sequence $<a_i>$ such that $a_1 Ea_0, a_2 Ea_1, \ldots, a_{n+1} Ea_n, \ldots$ all hold. Informally, it is sometimes stated that $E$ is well-founded just in case there are no “infinitely descending $E$-chains”. Once again, I submit that the notion of well-foundedness is both clear and unambiguous and, once again, the well-foundedness of a relation $E$ cannot be characterized in a first-order language (provided only that for each natural number $n$, there is a model of the background theory in which $E$ is well-founded and which contains $n + 1$ elements $a_0, \ldots, a_n$ such that $<a_1, a_0>, \ldots, <a_n, a_{n-1}>$ all satisfy $E$).

A second-order formulation is straightforward:

$$\forall X [\exists x Xx \rightarrow \exists x (Xx \& \forall y (Xy \rightarrow \neg yEx))].$$

---

11 Fix a first-order theory $T$ and model $M$ with domain $d$. Let $\Gamma$ be a set of formulas, containing a new predicate $P_x$, that purports to characterize the minimal closure of the extension of a formula $P$ under a unary function $p$. Under the following assumptions, there are models of $T + \Gamma$ in which the extension of $P$ is not the minimal closure of the extension of $P$ under $p$: (1) the given model $M$ can be extended to a model of $T + \Gamma$; (2) in each model of $T + \Gamma$, the extension of $P$ both contains the extension of $P$ and is closed under $p$; and (3) in the given model $M$, the minimal closure of the extension of $P$ under $p$ contains infinitely many members not in the extension of $P$. Actually, under these conditions, there are models of $T$ that can be extended to models of $T + \Gamma$ in more than one way and, thus (following the main result of Corcoran [1971]) $\Gamma$ fails to be a legitimate definition at all, let alone a definition of a minimal closure.
In arithmetic, proofs by induction and definitions by recursion presuppose that the predecessor relation and the "less-than" relation are both well-founded on the natural numbers. The major use of well-foundedness, of course, is in set theory where it is important that the membership relation be well-founded.

In [1980], Hilary Putnam argues against realism in set theory. In one section (pp. 468–469) he claims that the Löwenheim-Skolem theorems indicate that there is no "fact of the matter" concerning whether all sets are constructible or even whether a given countable set of real numbers is constructible. To support this, he introduces a theorem that for every countable set $s$ of real numbers, there is an $\omega$-model $M$ of set theory which contains $s$ and satisfies "every set is constructible". Of course, a realist will maintain that the given set $s$ may nevertheless be nonconstructible "in reality". Putnam replies:

But what on earth can this mean? It must mean, at the very least, that the model $[M]$ we have described [which contains $M$ and satisfies "all sets are constructible"] would not be the intended model. But why not? (p. 469).

He then goes on to argue that there are no grounds to claim that the model $M$ is "unintended" since $M$ satisfies all of the "theoretical constraints" that have been placed on intended model. The only "theoretical constraints" considered, however, involve the structure of the finite ordinals and the satisfaction of the axioms of first-order set theory. But one can surely claim that the well-foundedness of the membership relation is a "theoretical constraint" on (intended) models of set theory: one would hardly consider a structure with a non-well-founded membership relation to be an "intended model". Yet it follows (from a result that Putnam indicates) that if a given set $s$ is nonconstructible, then any model (containing $s$) that satisfies "$s$ is constructible" is not well-founded. Thus, against Putnam, in the above theorem, if $s$ is not constructible, then there is a clear sense in which $M$ is an unintended model of first-order set theory. I take these results as further evidence that a first-order language is not adequate to formalize set theory or, to borrow a phrase, to formulate the "theoretical constraints" on intended models of set theory.

2.2. In most formulations of arithmetic, the only second-order axiom is the statement of mathematical induction:

$$(I) \forall P([P0 \& \forall x(Px \rightarrow Psx)] \rightarrow \forall xPx).$$

In common formulations of real analysis, the only second-order axiom is that of completeness, the statement that every bounded subset of the domain has a least

---

12 A structure $M$ is an $\omega$-model of set theory iff the extension of "finite ordinal" under the relation of membership (in $M$) is isomorphic to the natural numbers under "less than".

13 It might be noted, as an aside, that even if the present considerations concerning higher-order languages are correct, they do not affect the bulk of Putnam's conclusions concerning reference in [1980] and [1981]. I suggest that it is not the Löwenheim-Skolem theorems that are relevant, but the fact that isomorphic structures satisfy the same set of sentences or, in other words, that a language—of any order—cannot distinguish among isomorphic structures.
upper bound:

\[(\text{C}) \quad \forall P(\exists x \forall y (P y \rightarrow y \leq x)) \rightarrow \exists x[\forall y(P y \rightarrow y \leq x) \& \forall z(\forall y(P y \rightarrow y \leq z) \rightarrow x \leq z)].\]

In second-order Zermelo-Fraenkel set theory (ZFC) the only second-order axiom is that of replacement, which states that for every function \(f\), the image of any (iterative) set under \(f\) is a set. Here, I give a formulation which contains a predicate (class) variable \(P\) (thought of as ranging over collections of ordered pairs), rather than a variable ranging over functions on the domain:

\[(\text{R}) \quad \forall P[\forall x \forall y \forall z(P(x, y) \& P(x, z) \rightarrow z = y) \rightarrow \forall x \exists y \forall z(z \in y \leftrightarrow \exists w (w \in x \& P(w, z))].\]

The usual first-order axiomatization of each of these theories is obtained by replacing the respective second-order axiom by a scheme. In each case, the second-order variable \(P\) is replaced by an “arbitrary” first-order formula \(\Phi(x)\), with \(x\) free. The result is an infinite number of axioms, one for each suitable formula \(\Phi\) of the respective first-order language:

\[(\text{I-}\Phi) \quad \Phi(0) \& \forall x(\Phi(x) \rightarrow \Phi(sx)) \rightarrow \forall x \Phi(x),\]

\[(\text{C-}\Phi) \quad \exists x \forall y(\Phi(y) \rightarrow y \leq x) \rightarrow \exists x[\forall y(\Phi(y) \rightarrow y \leq x) \& \forall z(\forall y(\Phi(y) \rightarrow y \leq z) \rightarrow x \leq z)],\]

\[(\text{R-}\Phi) \quad \forall x \forall y \forall z(\Phi(x, y) \& \Phi(x, z) \rightarrow z = y) \rightarrow \forall x \exists y \forall z(z \in y \leftrightarrow \exists w (w \in x \& \Phi(w, z))].\]

The difference between, say, second-order real analysis and first-order real analysis is that in the former it is asserted that the completeness property applies to every subset of the domain, whether it can be defined in the language of real analysis or not; whereas in the latter, it can only be shown that the completeness property applies to subsets of the domain that are definable in the given first-order language.

The purpose of this subsection is to argue that this restriction on the first-order theories is artificial—it does not conform to mathematical practice. This, of course, is not to deny the substantial utility of the metamathematical study of the first-order theories, but it is to deny that the first-order theories adequately express the mathematical practice of the respective fields. I begin with three considerations suggested by Kreisel [1967] concerning arithmetic and real analysis. This is followed by a discussion of second-order set theory.

2.2.1. The first consideration is epistemic. As indicated in §1 above, a basic presupposition of the present paper is that arithmetic and real analysis each has an intended interpretation independent of the language used to describe it. The theories in question are not taken as hypothetical or logistic systems. One can therefore inquire as to why a given axiom (or other statement) is believed or

---

14 The rich developments of nonstandard arithmetic and analysis are but two examples of the utility of the metamathematical study of first-order theories.
accepted, either by a particular mathematician or by the mathematical community as a whole. Kreisel wrote:

A moment’s reflection shows that the evidence of the first-order schema derives from the second-order [axiom]; the difference is that when one puts down the first-order schema, one is supposed to have convinced oneself that the specific formulae used ... are well-defined in any structure one considers. (Kreisel [1967, 148])

Kreisel’s point, I take it, is that a given mathematician believes or accepts the instances of the first-order scheme only because she (already) believes or accepts the second-order axiom. Suppose, for example, that one is given an instance of the completeness scheme (C-\Phi), perhaps one in which the indicated formula \Phi is rather complicated, and suppose that the person is asked whether it is true of the real number structure and, if it is, why she believes it. On the basis of the first-order axiomatization, the answer would be something like “I accept this formula because it is an axiom—it has the form of the completeness scheme”, or perhaps, “this formula is one of the basic or defining characteristics of the real number structure”. The present suggestion is that a more accurate, or (at any rate) more natural response would be: “The subformula \Phi(x) determines a set of real numbers (this follows from the comprehension axiom of second-order logic). The given formula (in the form (C-\Phi)) asserts that if this set is bounded, then it has a least upper bound. I accept this because it is an application of the completeness axiom—an axiom that characterizes the real number structure”.

As is well known, some philosophers hold that a given second-order language has undesirable or dubious commitments beyond those of its first-order counterpart. A like-minded mathematician might argue that by using first-order analysis, he is sacrificing epistemic clarity or simplicity for more acceptable ontological commitments. The onus on such a mathematician is to show why he accepts each instance of the scheme. A separate justification for each axiom is, of course, out of the question—there are infinitely many. Moreover, this mathematician cannot claim that he believes the instances of the scheme because they are the “safe” or “reasonable” consequences of the second-order axiom—the consequences which do not have the undesirable or dubious commitments. If one rejects the second-order axiom on ontological grounds, then one cannot use it as a premise to justify other sentences. A second possibility, perhaps, would be to formulate the completeness scheme informally as “every appropriately definable, bounded set of real numbers has a least upper bound”. As it stands, however, this statement involves quantification over sets of real numbers and, thus, is prima facie second-order. Moreover, it presupposes a concept of definability which probably is to be characterized as a minimal closure. I do not claim here that a plausible justification of the first-order scheme that does not involve a second-order language is impossible. I do suggest that such a justification has yet to be given.

2.2.2. Since the second-order axiomatizations of arithmetic, real analysis, and set theory do not contain schemes, the second-order theories are somewhat independent of the nonlogical terminology available in the language. For example, the characterization of the natural numbers in a second-order language containing only the constant 0 and a name for the successor function is essentially the same as the
characterization in a language containing names for other functions, such as addition and multiplication (the only difference being that the latter contains axioms to define those functions). This is not the case with the first-order versions. In arithmetic, the extent of the scheme \( I(\Phi) \) is determined by the available formulas \( \Phi \), and this, of course, depends on the nonlogical terminology of the formalizing language. Kreisel notes:

> The choice of the first-order schema is not uniquely determined by the second-order axiom! Thus, Peano's own axioms mention explicitly only the constant 0 and the successor function..., not addition nor multiplication. The first-order schema built up from 0 and [the successor function] is a weak ... subsystem of classical first-order arithmetic... and quite inadequate for formulating current informal arithmetic. (Kreisel [1967, 148])

Suppose, for example, that in the course of a treatise on the natural numbers, a mathematician decides to introduce a new function \( f \). She proceeds by adding a function letter, giving a description (e.g., a recursive derivation) of the function, and proving that a unique function is thereby described. The attitude is that the mathematician has introduced a new function on the same domain and, thus, that she is working in the same theory as she was before the function was introduced. A theorem (which may not mention the function \( f \)) in the "extended" theory is taken to be true of the natural numbers, even if it could not be deduced from the previous axioms alone.

This attitude is reflected in the second-order axiomatization of arithmetic. In this case, the introduction of the function \( f \) does not alter the basic description of the natural numbers. Moreover, the indicated proof that a unique function has been introduced amounts to a demonstration of what may be called "unique extendibility"—a demonstration that each model of the original axiomatization can be extended to a model of the new theory in exactly one way. It follows from the main result of Corcoran [1971] that the characterization of \( f \) is semantically eliminable and noncreative or, in other words, that the requirements of an acceptable definition have been met. In short, in the second-order theory, all is as it should be.

This is not the case with first-order axiomatizations. In this case, the introduction of a new function letter extends the language and, thus, extends the set of formulas in the form \( \Phi(x) \). It follows that the induction scheme \( I(\Phi) \) is itself extended. That is, the introduction of a new function letter results in a change in the basic description or axiomatization of the original theory. It is as if one is working in a new theory. Moreover, the new theory may not even have the "same" models as the original. To elaborate Kreisel's example, first-order Peano arithmetic formulated with only 0 and the successor function is not uniquely extendible vis-à-vis addition. There are models of the original theory that cannot be extended to models of arithmetic-with-addition, and there are models of the original that can be so extended in more than one way. To reapply the result of Corcoran [1971], it follows that the introduction

---

15 In the second-order case, if the new function \( f \) is defined by primitive recursion, then the eliminability and noncreativeness can be proven within the standard metatheory of the axiomatization.
of addition to first-order Peano arithmetic cannot be accomplished with a legitimate definition. (It follows, incidentally, that in the case at hand, the "informal" proof that a unique function is characterized cannot be formulated in a first-order language.)

2.2.3. With the use of schemes which depend on the language of the theory, first-order arithmetic and first-order analysis are presented as isolated theories which are independent of both each other and the rest of mathematics. This is not in accord with the practice of viewing such mathematical structures as interrelated. This practice is manifest in the common technique of "embedding" or "modeling" one structure in another. In Kreisel's words:

... very often the mathematical properties of a domain \( D \) become only graspable when one embeds \( D \) in a larger domain \( D' \). Examples: (1) \( D \) integers, \( D' \) complex plane; use of analytic number theory. (2) \( D \) integers, \( D' \) \( p \)-adic numbers; use of \( p \)-adic analysis. (3) \( D \) surface of a sphere, \( D' \) 3-dimensional space; use of 3-dimensional geometry. Non-standard analysis [also applies] here ... (Kreisel [1967, 166])

To take a simple example, when one realizes that the set-theoretic hierarchy (or, for that matter, the complex plane) contains the natural numbers—or an "isomorphic copy" of the natural numbers—then one can use set theory (or complex analysis) to shed light on the natural numbers. That is, since isomorphic structures have the same set of truths, a theorem of set theory that refers only to the "natural-numbers-of-set-theory" is true of the natural numbers.

It is well known that this technique can produce results that are not obtainable in the original theories. In the indicated example, this happens because there are subsets of the natural numbers which are definable in set theory, but are not definable in arithmetic alone. Thus, one who accepts first-order arithmetic as adequate cannot make the straightforward claim that some theorems of set theory reflect truths of the natural numbers. Indeed, many of the indicated arithmetic statements are false in some models of first-order arithmetic.

The practice of embedding structures indicates a further area in which categoricity is important. The point is that in order to embed a structure \( D \) into a structure \( E \), one must have a means of recognizing a substructure of \( E \) as isomorphic to \( D \). Otherwise, one cannot be certain that \( D \) really is a substructure of \( E \). The formal analogue of this requirement is a categorical characterization of \( D \). Suppose, for example, that someone believes that a structure \( M \) "contains" the natural numbers and, thus, that the study of \( M \) may produce (new) theorems of arithmetic. In attempting to verify this, he defines a certain substructure \( N \) of \( M \) and shows that this structure satisfies the axioms of first-order arithmetic. Since the latter is not categorical, the mathematician cannot conclude that \( N \) is isomorphic to the natural numbers, nor can he conclude that all theorems of \( M \) whose quantifiers are restricted to the domain of \( N \) are true of the natural numbers. For all he knows (so far), the substructure \( N \) may be a nonstandard model of arithmetic. The situation is perhaps analogous to an observation that a certain set-theoretic structure is a group. Since the group axioms are not categorical, it does not follow that the properties of this structure are "truths" of group theory. In the example at hand, of course, the situation would be different if the mathematician showed that \( N \) satisfies the axioms
of second-order arithmetic. In this case, he can conclude that \( N \) is isomorphic to the natural numbers and, thus, that any theorem of \( M \) whose quantifiers are restricted to the domain of \( N \) is true of the natural numbers.

2.2.4. This section concludes with some remarks on second-order set theory (ZF). Of course, the intended interpretation of ZFC is not itself an iterative set: \(^{16}\) the set-theoretic hierarchy is not a member of itself. Thus, one who believes that all legitimate collections are (isomorphic to) sets may balk at the range of the second-order variables and, consequently, may have trouble envisioning second-order ZFC as a theory about the set-theoretic hierarchy \(^{17}\) (see Boolos [1975]). Of course, even the totality of the (intended) range of the variables of first-order set theory is not a set, but at least every element thereof—every element referred to by the theory—is a set.

It might be noted that it is common for set-theorists to speak of the set-theoretic hierarchy itself, at least in informal language. For example, for a given formula \( \Phi(x) \) and structure \( M \) with domain \( d \), it is stated that \( \Phi \) is absolute in \( M \) just in case for each \( x \in d \), \( M \models \Phi(x) \) iff \( \Phi(x) \) holds in the set-theoretic hierarchy.

Moreover, there is at least no formal antinomy involved in using second-order ZFC to describe the set-theoretic hierarchy. Indeed, second-order ZFC is deductively equivalent to the so-called Morse-Kelley set theory (MK), a first-order theory with two variable sorts, one ranging over sets, the other over “classes”. It follows from a well-known theorem concerning MK that second-order ZFC is consistent if the theory consisting of first-order ZFC and an axiom asserting the existence of one inaccessible cardinal is consistent. Following the theme of this article, however, I suggest that what makes the second-order version attractive is not its deductive strength, \(^{18}\) but rather its semantics.

The considerations of the previous ?? 2.2.1-2.2.3 apply to set theory, but only to a limited extent. The best case can be made for the epistemic point of ?2.2.1: one accepts the instances of the first-order scheme of replacement (R-\( \Phi \)) only because one accepts the second-order version (R). The possibility of introducing new terminology, which would extend the scheme (R-\( \Phi \)) is moot (at present)—I do not

\(^{16}\) To follow footnote 10, in this subsection (only) the word “set” is taken as “iterative set” or, in other words, as “member of the set-theoretic hierarchy”. Also, the terms “model” and “interpretation” refer to structures whose domains are (isomorphic to) iterative sets. As above, in the context of set theory, logical sets are designated “classes”.

\(^{17}\) It might be noted, as an aside, that when one envisions a (set-theoretic) model \( M \) of set theory, these “commitments” are unexceptionable. The domain \( d \) of such a model is itself a set. The statement that \( M \) satisfies the second-order replacement axiom (R) contains variables ranging over subdomains of \( d \). Such items are themselves sets. Thus, the statement that “\( M \) is a model of second-order ZFC” only requires first-order variables.

\(^{18}\) There are, perhaps, a few items of interest concerning the relative deductive strength of second-order ZFC (i.e. MK) over first-order ZFC. Let \( T \) be the collection of “set” theorems of second-order ZFC: \( \Phi \in T \) iff \( \Phi \) is a first-order sentence provable in second-order ZFC. It is easily seen that \( T \) contains sentences not provable in first-order ZFC. Moreover, the “natural” first-order extensions of first-order ZFC whose theorems contain \( T \) are much stronger (vis-à-vis relative consistency) than second-order ZFC. Second, one can prove in second-order ZFC that there is a countable set \( d \) which is first-order elementarily equivalent to the set-theoretic hierarchy. This theorem represents a rather natural (informal) application of the downward Löwenheim-Skolem theorem. Yet the result in question cannot be stated, much less proved, in first-order ZFC.
know of any relevant cases. Concerning the embedding of theories, one can perhaps view the construction of Boolean-valued models as an embedding of the set-theoretic hierarchy in a richer structure, but, of course, Boolean-valued models can be reinterpreted in the set-theoretic hierarchy.

As noted above, it is essential to the concept of set that the membership relation be well-founded. It is useful at this point to look at the axiom that "asserts" this, the axiom of Foundation. It is a first-order sentence:

\[(F) \forall x (\exists y (y \in x) \rightarrow \exists y (y \in x \land y \cap x = \emptyset)).\]

The axiom (F) asserts that every nonempty set has a member that is disjoint from it. In first-order set theory, this axiom does prevent definable, infinitely descending \(\omega\)-chains (such as finite, closed \(\omega\)-sequences), but, as noted above, neither (F) nor any first-order improvement precludes non-well-founded models. The important point here is that in second-order \(\text{ZFC}\) the same axiom (F) does insure the well-foundedness of membership. That is, the axiom of foundation, a first-order sentence, only "works" in second-order set theory.

More can be said. There is a well-known theorem that a structure \(M = \langle d, E \rangle\) is a model of second-order \(\text{ZFC}\) iff there is an inaccessible cardinal \(\kappa\) such that \(M\) is isomorphic to \(V\kappa\).\(^\text{19}\) Thus, up to isomorphism, the models of second-order \(\text{ZFC}\) are certain "initial segments" of the set-theoretic hierarchy. It may be concluded that the second-order theory characterizes the set-theoretic universe in every respect except the "size" of the class of ordinals. Because of this, Weston [1976] calls second-order set theory "almost categorical" (although he denies that this is significant).

Of course, any model of second-order \(\text{ZFC}\) is also a model of first-order \(\text{ZFC}\). I suggest that if \(M = \langle d, E \rangle\) is any other model of first-order \(\text{ZFC}\)—that is, if \(M\) is not isomorphic to an inaccessible rank—then there is a clear sense in which \(M\) is nonstandard. There are three possibilities. First the relation \(E\) may not be well-founded. This, by itself, would rule out \(M\) as standard. If, on the other hand, \(E\) is well-founded, then \(M\) is isomorphic to a transitive set. Thus, \(E\) may be taken as the membership relation on transitive \(d\). In this case, a second possibility is that \(M\) is not closed under the subsets of its elements. That is, there are sets \(b \in d\) and \(a \subseteq b\) such that \(a \notin d\). I take it that this also disqualifies \(M\)—standard models should contain all of the subsets of each element, whether they are definable or not. The third possibility is that \(M\) fails to satisfy the replacement axiom, even on its ordinals: There is an ordinal \(\gamma\) in \(d\) and a sequence of \((\gamma\text{-many})\) ordinals \(\langle a_{\beta} \rangle_{\beta \in \gamma}\) such that sup \(\alpha\) is not in \(d\). I submit that this also disqualifies \(M\) as a standard model.

\S 3. Other languages. Previous sections have been devoted to the claim that axiomatizing the various branches of mathematics with separate first-order theories does not reflect important aspects of mathematical practice. One may be convinced

\(^{19}\) The usual definition of "\(\kappa\) is inaccessible" is "\(\kappa\) is regular and for every \(\alpha \in \kappa\), \(2^\alpha \in \kappa\)." Although in practice this is a rather useful definition, an equivalent one that perhaps illustrates the "inaccessibleness" would be "\(V\kappa\) is a model of second-order \(\text{ZFC}\)." Informally, the former definition indicates that \(\kappa\) cannot be "obtained" from smaller ordinals by the operations of powerset and \(\alpha\)-fold union (for any \(\alpha \in \kappa\)); the latter definition, that \(\kappa\) cannot be "obtained" from any sets of smaller rank by any operation or function implied by the axioms of second-order \(\text{ZFC}\).
of this, of course, without believing that it is necessary to provide second-order theories. This section examines several alternatives. §3.1 concerns three “intermediate” languages—infinitary languages, \( \omega \)-languages, and free-variable versions of second-order languages. The conclusion is that, among these, only the latter substantially overcomes the deficiencies of first-order languages. §3.2 concerns the program of using the language of first-order set theory to axiomatize the various branches of mathematics.

3.1. It is easily seen that most of the above considerations against first-order languages apply to any language whose semantics is compact. Thus, this subsection is limited to languages with noncompact semantics. Of course, in such cases, any “sound” deductive system is not complete: there are consistent sets of sentences that are not satisfiable.

3.1.1. Infinitary languages are easily dismissed. There is little doubt that the study of such languages has proven to be a fruitful and insightful branch of mathematical logic, but it need hardly be mentioned that infinitary languages are not serious candidates for the underlying language of mathematics. One of the chief purposes of language is to facilitate communication. Minimally, to be successful for communication, a given sentence must be capable of being spoken or written in a finite amount of time, using a finite amount of materials, etc.

Of course, in the language of informal mathematics, it is possible to describe the formation rules and some of the formulas of an infinitary language (such as infinite conjunctions and disjunctions). Thus, one may claim that the language of mathematical practice is best formulated as a (finitary) metalanguage for an infinitary object language whose subject matter is one of the various structures. Notice, however, that since the intended interpretation of the metalanguage in question is itself an infinite structure—an infinitary language—many of the above considerations apply. In short, the conceived metalanguage cannot be first-order.

3.1.2. More serious candidates, perhaps, are languages that allow quantification over the (standard) natural numbers: an \( \omega \)-language is a language that contains two variable sorts, one of which ranges over the intended domain, the other over the natural numbers (see Barwise [1977, 42–44]). It is required that in every interpretation, the range of the “natural number variables” be isomorphic to the natural numbers. In such languages, functions and relations involving the natural numbers and the intended domain can be introduced by primitive recursion. In what follows, let \( m, n, \ldots \) be variables ranging over the natural numbers and \( x, y, \ldots \) be variables ranging over the intended domain.

To begin with, first-order \( \omega \)-languages do not have all of the shortcomings of first-order languages discussed in §2.1 above. In particular, such languages can characterize individual minimal closures. For example, if \( \Phi(x) \) is a formula and \( p \) denotes a unary function on the domain, one can introduce a relation \( R(m, x) \) between natural numbers and the domain by primitive recursion as follows:

\[
R(0, x) \iff \Phi(x); \quad R(sn, x) \iff \exists y(R(n, y) \& x = py).
\]

---

20 Such an account was developed by Zermelo [1931] (see also Moore [1980, 124–127]). For Zermelo, the “metalanguage” for this infinitary object language is, in effect, a second-order set theory.
The minimal closure of the extension of \( \Phi \) under the function denoted by \( p \) is then characterized by the formula \( \exists n R(n, x) \).

Well-foundedness, however, cannot be characterized by a first-order \( \omega \)-language. This is indicated by the fact that there are \( \omega \)-models of first-order ZFC that are not well-founded. It is easily seen that such structures are also models of virtually any version of ZFC formulated in a first-order \( \omega \)-language.

Finally, for any given formula \( \Phi(x) \) with only \( x \) free, it can be stated that the extension of \( \Phi \) is finite. If a new function letter \( f \) (from the natural numbers to the domain) is introduced, then the sentence:

\[
\forall n \left( \exists x (\Phi(x) \& \forall m < n (x \neq fm)) \rightarrow \Phi(fn) \& \forall m < n (fm \neq fn) \right) \\
\land \exists n (\neg \Phi(fn) \lor \exists m < n (fn = fm))
\]

is satisfied by all those, and only those, interpretations of the rest of the theory in which the extension of \( \Phi \) is finite.

Similarly, if \( g \) and \( h \) are new function letters, then the sentences

\[
\forall n \forall m (\Phi(gn) \& (m \neq n \rightarrow gn \neq gm)) \quad \text{and} \quad \forall x (\Phi(x) \rightarrow \exists n (hn = x))
\]

are satisfied by all those, and only those, interpretations of the rest of the theory in which the extension of \( \Phi \) is, respectively, infinite and countable.

This, however, is the limit of the ability of first-order \( \omega \)-languages to characterize the cardinality of an extension. Indeed, the semantic version of the downward Löwenheim-Skolem theorem applies: For a (countable) \( \omega \)-language, any structure whose domain is infinite has a countable substructure that is elementarily equivalent to it.

On the other hand, the upward version of the Löwenheim-Skolem theorem does not apply to first-order \( \omega \)-languages. It is possible, in particular, to provide categorical axiomatizations of some countable structures. Of course, one can easily characterize the natural numbers (up to isomorphism) with a first-order \( \omega \)-language, but if anything has the advantages of theft over toil, this does. Less trivially, the rational numbers can be characterized as an infinite field whose universe is the minimal closure of \( \{1\} \) under the field functions and their inverses.

Moreover, even the theory of real analysis formulated in a first-order \( \omega \)-language is an improvement over the usual first-order version. In the former, one can apply the completeness scheme to enough sets defined as minimal closures to insure that all the models are Archimedean. It follows that each such model is isomorphic to a subset of the real numbers.\(^{21}\)

\(^{21}\) Another group of languages that may be considered are \( \omega \)-languages that have variables ranging over functions from the natural numbers to the intended domain. For infinite domains, such languages are equivalent to those with variables ranging over countable subsets of the domain. In this context, the above "partial" characterization of finitude can be extended to a full characterization, and well-foundedness can be straightforwardly characterized. Moreover, in real analysis, the completeness scheme would imply that every bounded, countable subset of the domain has a least upper bound. This, of course, entails that the characterization is categorical. Finally, even the characterization of set theory is a significant improvement in the sense that every model thereof is well-founded and closed under its countable subsets.
From the present point of view, the major shortcoming of \( \omega \)-languages is that they assume or presuppose the natural numbers. Therefore, such a language cannot be used to show, illustrate, or characterize how the natural number structure is itself understood, grasped, or communicated. The attractiveness of such languages is perhaps that, by studying their model theory, one can learn which structures can be characterized in terms of the natural numbers. Thus, such a language might be useful to someone who accepts a categorical axiomatization of arithmetic, but is skeptical of the characterizations of other, richer domains. (Analogous remarks apply to the so-called \( M \)-languages, where \( M \) is any fixed structure.)\(^{22}\)

3.1.3. The final “intermediate” language considered is one that contains free, but not bound, second-order variables. Such a language \( L \) can be obtained from a first-order language by adding a list \( X, Y, \ldots \) of (free) predicate variables and a list \( f, g, \ldots \) of (free) unary function variables (and modifying the formation rules accordingly). A formula of the extended language \( L \) is called a sentence if it has no free first-order variables. If \( \Phi(x_1, \ldots, x_n, f_1, \ldots, f_m) \) is such a sentence all of whose second-order variables are indicated, and \( M \) an interpretation with domain \( d \), then \( M \models \Phi \) just in case \( M \) satisfies \( \Phi \) under every assignment of subsets of \( d \) to the variables \( X_i \) and functions on \( d \) to the variables \( f_j \). Thus, \( \Phi \) is treated as a universally quantified formula (equivalent to \( \forall x_1 \cdots \forall x_n \forall f_1 \cdots \forall f_m \Phi \)). We call such languages diminished second-order languages.

The diminished second-order languages differ from full second-order languages in several ways: (1) The only second-order quantification permitted in the former is, in effect, prenex universal quantification. That is, instead of arbitrarily many quantifiers arbitrarily embedded, the diminished versions allow only universal quantifiers whose scope is the entire formula.\(^{23}\) (2) The diminished second-order languages have neither \( n \)-ary function variables nor \( n \)-ary relation variables for any \( n > 1 \). (3) Diminished second-order languages have no higher-order predicate, relation, or function constants.

Similar languages are explicitly formulated and studied in Corcoran [1980, 192ff.] as “slightly augmented first-order languages” (the difference being that Corcoran’s languages have only a single free predicate variable and no function variables). Church [1956, §48] implicitly classifies diminished second-order languages as “applied” first-order languages.

It is easily seen that virtually none of the above shortcomings of first-order axiomatizations are shared by diminished second-order axiomatizations. The above

\(^{22}\) It might be noted that there is a sense in which the expressive power of first-order \( \omega \)-languages is “equivalent” to that of first-order languages augmented with an “ancestral operator”: If \( Rxy \) is a binary relation (such as parenthood) we say that \( b \) is an \( R \)-ancestor of \( c \) just in case there is a finite sequence \( a_0, \ldots, a_n \) such that \( a_0 = c, a_n = b \) and \( Ra_{i+1}a_i \) for each \( i, 0 \leq i \leq n \). An ancestral operator is a variable-binding operator \( \Gamma \) such that for each formula \( \Phi(x, y) \) with \( x \) and \( y \) free, \( (xy\Gamma)z \Phi \) \( zw \) is a formula equivalent to “\( w \) is an ancestor of \( z \) under the extension of \( \Phi^\omega \).” By a construction similar to that for minimal closures, an ancestral operator can be formulated in an \( \omega \)-language. Conversely, the natural numbers can be characterized up to isomorphism in a first-order language augmented with an ancestral operator.

\(^{23}\) A sentence \( \Phi(X) \) of \( L \) is equivalent to the second-order “every subset of the domain satisfies \( \Phi \).” In general, there is no sentence of \( L \) that is the contradictory of this. The “contrary” \( \neg \Phi(X) \) is equivalent to “no subset of the domain satisfies \( \Phi \).”
characterizations of minimal closure, finitude, and well-foundedness, as well as the standard axiomatizations of arithmetic, real analysis, and set theory involve only prenex universal quantification. In these cases, at least, no other higher-order quantification is needed.\(^{24}\)

In the present article, no stand is taken on the issue of whether diminished (or even full) second-order languages are sufficient to axiomatize branches of mathematics.\(^{25}\)

The only claim made is that some second-order variables are necessary. In short, the present thesis is that diminished second-order languages serve as a "lower bound" on languages to formulate mathematical theories, and not necessarily a "greatest lower bound".

3.2. Perhaps another alternative to second-order languages would be to formulate mathematical theories in the language of first-order set theory. In general, with any theory \(T\) formulated in a second-order language by a finite number of axioms, there corresponds (in a straightforward manner) a formula \(T(x)\), of first-order set theory, which amounts to "\(x\) is a model of \(T\)". Every sentence \(\Phi\) of the language of \(T\) can then be "translated" into a sentence \(\Phi' = \forall x(T(x) \rightarrow \Phi_x)\) in the language of set theory, where \(\Phi_x\) is obtained from \(\Phi\) by replacing all first-order variables by (set) variables ranging over (i.e., relativized to) the "domain" of \(x\), replacing all predicate variables by variables ranging over the subsets of the domain of \(x\), etc.

Thus, if \(\Phi\) is a sentence in the language of \(T\), then the set-theoretic \(\Phi'\) amounts to "\(\Phi\) is true in all (set-theoretic) models of \(T\)". Since the set-theoretic hierarchy is usually taken to provide the semantics of second-order languages, \(\Phi'\) amounts to "\(\Phi\) is a semantic consequence of \(T\)". Also, under normal conditions concerning the relative strengths of deductive systems, a proof of \(\Phi\) in \(T\) can be routinely translated into a proof of \(\Phi'\) in set theory. That is, concerning object language proofs, first-order set theory can do anything a second-order language can do, usually more.\(^{26}\)

It might even be added that the concepts and properties discussed in §2.1 above have straightforward characterizations in (first-order) set theory. For example, if \(b\) is a set and \(c\) is a set of functions, then "\(x\) is in the minimal closure of \(b\) under the members of \(c\)" is characterized by the following formula:

\[
\text{MC}(b, c, x) : \forall y([b \subseteq y \land \forall z \forall s \forall w (z \in y \land s \in c \land \langle z, w \rangle \in s \rightarrow w \in y)] \rightarrow x \in y).
\]

\(^{24}\) Actually, for the theories whose domains are infinite, the function variables may be eliminated by introducing constants for pairing and unpairing functions. In such cases, predicate variables can "play the role" of function variables.

\(^{25}\) Kreisel [1967] seems to suggest that some third-order axiomatizations may be required. Some writers seem to envision \(n\)th order languages, where \(n\) is any ordinal. There is also an issue as to whether the deductive strength of diminished second-order languages is adequate for mathematical practice.

\(^{26}\) It might be added that many metalanguage statements can be formulated in first-order set theory. For example, the sentence \(\exists x T(x)\) amounts to "\(T\) is satisfiable". The categoricity of \(T\) is asserted by a set-theoretic formula \(C(T)\) of the form \(\forall x \forall y (T(x) \land T(y) \rightarrow (x \land y \text{ are isomorphic}))\); if \(T\) has been proven categorical, then — up to the relative strength of set theory over the metatheory of \(T\)—the sentence \(C(T)\) is a theorem of set theory.
Notice that \( \forall b \forall c \exists ! y \forall x (x \in y \iff MC(b, c, x)) \) is a theorem of set theory. That is, it is provable that for every \( b \) and \( c \) there is a unique minimal closure of \( b \) under \( c \).

Thus, it might seem that first-order languages have been revived. I submit, however, that any formulation of the various branches of mathematics in first-order set theory does not reflect mathematical practice. It is not sufficient for an axiomatization to get the appropriate theorems; the semantics must also be correct.

Different reasons are given for this negative judgement, depending on whether the background language of first-order set theory is taken as an interpreted language or an uninterpreted language. In short, if the language is considered to be uninterpreted, then the above advantages are, in a certain sense, merely formal and illusory: they do not apply to the overall semantics. If the language of set theory is considered to be interpreted, then the advantages are those of theft over toil. Indeed, if \( M \) is the intended interpretation, then the question remains as to how \( M \) is itself grasped, understood, and communicated. Moreover, there is a clear sense in which the presuppositions of an interpreted set theory are greater than those of the semantics of second-order languages.

Consider first the case in which the background language of set theory is taken as uninterpreted. Since this language is first-order, it has nonisomorphic models, some of which are clearly nonstandard. There are models, for example, in which the membership relation is not well-founded. This observation, in effect, undermines the above “advantages” to the present program.

To focus on an example, consider the formula \( N(x) \) of set theory that asserts that “\( x \) is a model of the natural numbers”. The fact that second-order arithmetic is known to be categorical corresponds to a set-theoretic theorem of the form

\[
\forall x \forall y (N(x) \land N(y) \rightarrow (x \text{ and } y \text{ are isomorphic})).
\]

It follows that for each model \( M \) of the background set theory, if \( a \) and \( b \) are in the domain of \( M \) and \( M \models N(a) \) and \( M \models N(b) \), then \( M \models (a \text{ and } b \text{ are isomorphic}) \). In other words, the above theorem entails that within the same model of set theory, any two sets satisfying \( N(x) \) are isomorphic (in that model). This is not enough. The preformal understanding of the categoricity of arithmetic is that any two models of arithmetic are isomorphic, not just any two within the same model of set theory. By a straightforward compactness argument, it is easy to see that if the set theory is consistent, then it has a model \( M' \) with an element \( b \) (of the domain thereof), such that \( M' \models N(b) \), but the collection of \( M' \)-elements of \( b \) is not isomorphic to the natural numbers. This, of course, is a variant of the Skolem “paradox”. Similar considerations apply to any theory that has an infinite model.

Virtually the same considerations apply to the (first-order) set-theoretic versions of the items of §2.1. Consider, for example, the formula \( MC(b, c, x) \) corresponding to “\( x \) is in the minimal closure of \( b \) under \( c \)”. There is a model \( M' \) of set theory containing elements \( b, c, d \) such that \( M' \models \forall x (x \in d \iff MC(b, c, x)) \), but the collection of \( M' \)-elements of \( d \) is not the required minimal closure.

Thus, the practice of formulating mathematical theories in an uninterpreted first-order set theory does not preclude nonstandard or unintended interpretations. Such interpretations occur in nonstandard or unintended interpretations of the back-
ground set theory. Once again, I agree with Skolem that the result is an unavoidable relativity of all mathematical notions. However, I submit that such a relativity does not reflect mathematical practice.27

I turn to the case in which the (first-order) background language of set theory is taken as interpreted. The background theory need not be ZFC, of course; it may be a formalized version of informal set theory. Let $M$ be the intended interpretation. Presumably, $M$ is standard in the sense that it is well-founded, extensional, and closed under the subsets of its elements. For simplicity of treatment, assume that $M$ does not contain urelements. It follows that $M$ is isomorphic to a limit rank $V\lambda$. If $M$ is to be adequate for arithmetic, geometry, analysis, functional analysis, etc., the ordinal $\lambda$ must be at least $2\omega$—otherwise, there are no models of the above theories in $M$. If $M$ is to be adequate for ZFC (and perhaps category theory), then $\lambda$ must be larger than an inaccessible cardinal.

Notice, first, that in the present situation, the considerations against uninterpreted languages do not apply. For example, the formula $N(x)$ is interpreted as “$x$ is a set in $M$ that is a model of the natural numbers”. Since $M$ is standard, all such sets $x$ are isomorphic to the natural number structure.

Recall that from the present perspective, the major shortcoming of the $\omega$-languages is that they presuppose an understanding of the natural numbers. A similar, and perhaps more serious, problem is found in the present use of interpreted set theory. To say that a structure $P$ is characterized up to isomorphism by the language of set theory as interpreted is only to say that $P$ can be characterized in terms of $M$ or “up to $M$”. The problem as to how $M$ is itself grasped, understood, or communicated is left open. Moreover, I suggest that this latter problem is more difficult than the original problem of accounting for how the natural number structure, the real number structure, etc. are grasped, understood, or communicated. That is, the present program is a case of reducing one problem to a more difficult one. Without an independent characterization of $M$, it is not clear how the language of set theory overcomes the problems of characterizing structures in first-order languages.

At this point, perhaps, it might be suggested that one need not actually characterize the structure $M$. It is sufficient to let $M$ be any fixed model of the background set theory. It needs to be pointed out, however, that not just any model will do. To correctly characterize the requisite structures and concepts, $M$ must be a standard model of set theory. At a minimum, $M$ should be well-founded and closed under the subsets of its elements. As above, these are not first-order concepts.

It is instructive to compare the presuppositions of a given theory as formulated in a second-order language with those of the same theory as formulated in an interpreted language of set theory. The usual examples of arithmetic, real analysis, and set theory are considered.

---

27 Notice that none of these considerations apply if the uninterpreted background language of set theory is second-order (considered, as usual, with standard semantics). For example, if $M$ is a model of second-order set theory and $M \models N(b)$, then the collection of $M$-elements of $b$ is in fact isomorphic to the natural numbers under the indicated successor relation. This holds even though $N(x)$ is a first-order formula.
It seems safe to say that a theory formulated in a second-order language presupposes at least one model. Formalism and logicism aside, if there is no model of a theory (that is, if the theory is not satisfiable), then it has no possible subject matter. It is hard to imagine someone devoting time to studying a theory if he did not believe that it has an interpretation. Thus, arithmetic presupposes a denumerable structure with a successor function, real analysis presupposes a complete ordered field of cardinality $2^{\aleph_0}$, and set theory presupposes a hierarchy of (at least) inaccessible cardinality.

I follow the Quinean view that the ontological commitments of a theory are the values of its variables. Thus, a second-order version of a theory $T$ has presuppositions beyond those of its first-order counterpart. Indeed, the former has more variables. If $Q$ is an intended model of $T$, and $q$ the domain of $Q$, then the second-order theory presupposes the existence of each element of $q$, each subset of $q$, each function from $q$ to $q$, and each relation on $q$. Since the domain of $Q$, together with its subsets, functions, and relations, exhausts the variable-ranges of the second-order formulation of $T$, it follows that this list exhausts the indicated ontological presuppositions.

The important point here is that the second-order formulation of $T$ only presupposes the subsets of, and functions and relations on, a domain that is already presupposed. Indeed, the elements of $q$ are presupposed by the acceptance of any version of $T$: these elements are the values of the first-order variables. In particular, note that a second-order theory short of set theory does not presuppose a set-theoretic hierarchy.

The presuppositions of the formulation of $T$ in an interpreted language of set theory are thus much greater than the presuppositions of the formulation of $T$ in a second-order language (or, for that matter, the formulation of $T$ in an $n$th order language, for virtually any $n$). As above, let $M$ be the given interpretation of the background set theory and let $d$ be the domain of $M$. Presumably, each element of $d$ is presupposed by the acceptance of the background language as interpreted. If the standard models of $T$ are infinite, then to be adequate for $T$, the domain $d$ must contain an infinite set $c$. Presumably, the background theory has a powerset axiom (or theorem). Thus, the overall commitments of the set-theoretic program include the powerset of $c$, the powerset of the powerset of $c$, etc. In many set theories, this

---

28 It might be noted that the same presupposition applies to a theory formulated in a first-order language, and, in this case, it could be added that a standard model is presupposed. Suppose, for example, that it should happen that a dramatic event causes that mathematical community to believe that although first-order analysis is consistent, it does not have a standard model. (Suppose, for example, that the community comes to believe that every model of the theory has an undefinable, bounded subset that does not have a least upper bound.) Of course, the consistency of the first-order theory implies that it is satisfiable. In fact, the theory could be reinterpreted as a complicated structure of natural numbers. I suggest, however, that this fact would give little comfort to one who has devoted her life to studying a nonexistent real number structure.

29 The principle of accepting collections (as opposed to functions and relations) of previously accepted entities is held by such nominalists as Nelson Goodman [1972] and Hartry Field [1980]. For these philosophers, the objection to mathematical theories lies with the acceptance of the original first-order versions.
process can be iterated into the transfinite. In any case, these presuppositions are much more than one needs for arithmetic, real analysis, or just about any theory short of set theory.

Of course, it is to be conceded that despite the greater presuppositions, there are advantages to using a background language of set theory to formulate mathematical theories. Probably the most important of these is that the set-theoretic program provides a single, uniform "foundation" for all (or most) of mathematics. Indeed, as stated above, it is common to take mathematical theories and structures to be interrelated. This presupposes a common semantics for different branches and, hence, may indicate the propriety of a common language—a language of set theory. Consider, for example, the technique of embedding one structure, such as the natural numbers, into another, the complex plane. In the present framework, the embedding has a straightforward account in the set-theoretic object-language common to the two theories. It amounts to a theorem that for any model of the natural numbers and any model of the complex plane, there is a one-to-one homomorphism from the former into the latter.

With this uniform foundation, however, comes a uniform group of presuppositions—the intended model $M$ of the background set theory. This, I submit, is unnatural. Informally at least, the presuppositions of arithmetic are less than those of real analysis. The former has the relatively modest commitment to a denumerable set (plus, perhaps, its subsets, etc.) while the latter is committed to an uncountable continuum (plus, perhaps, its subsets, etc.). One who is doing arithmetic alone does not appear to be committed to, say, the real number line (or the set-theoretic hierarchy). From this perspective, a decision to use real analysis to study the natural numbers involves as expanded commitment.\(^{30}\) Kreisel [1967] observes that the desire to keep track of presuppositions is implicit in the practice of at least some mathematicians and, moreover, that this desire is reflected in the program of providing separate (second-order) theories of each branch:

... Bourbaki [for example] is extremely careful to isolate the assumptions of a mathematical theorem, but never the axioms of set theory implicit in a particular deduction ... This practice is quite consistent with the assumption that what one has in mind when following Bourbaki's proofs is the second-order axioms ... (Kreisel [1967, 151])

It is agreed, of course, that there is a common semantics for the various second-order languages. It does not follow, however, that there are substantial presuppositions and commitments for this semantics that apply uniformly to any theory

\(^{30}\) As indicated in footnote 26, a related "advantage" of the set-theoretic program is illustrated by the fact that for a given theory $T$, the "object language" statements of $T$ are interpreted as statements about the models of $T$ or, in effect, as statements about the semantics of $T$. Thus, the set-theoretic program provides a clear and straightforward connection between a theory and its semantics. It follows from Tarski's theorem, however, that, in some sense, the metatheory or semantics of a branch of mathematics is stronger (or, at any rate, substantially different) than the object language theory of that branch. Once again, for the purpose of keeping track of presuppositions, it is worthwhile to keep theory and metatheory separate.
formulated in a second-order language. Notice that in any case, the common metatheory, or semantics, of second-order languages in weaker than standard set theory, say ZFC. As observed by Boolos [1975], there is a straightforward “interpretation” of second-order logic in (first-order) set theory. That is, for each (effectively presented) second-order language $L$, there is a formula $\Theta(n)$ in the language of first-order set theory, such that for each natural number $n$, $\Theta(n)$ is true (in the set-theoretic hierarchy) just in case $n$ is the Gödel number of a valid sentence of $L$. From Tarski’s theorem, it follows that there is no set-theoretically definable translation $T$ from the language of set theory to $L$ such that for each sentence $\Phi$ of set theory, $\Phi$ is true iff $T(\Phi)$ is valid.

§4. First-order logic and second-order logic. The main conclusion of this article is that an adequate formalization of such branches of mathematics as arithmetic, real analysis, and set theory must involve (at least) a second-order language. I suggest, then, that the natural underlying logic of these branches is (at least) second-order.

As indicated above (see footnote 2), there has been some work in recent years aimed at showing that second-order logic is not logic at all. Such arguments usually focus on either the ontological presuppositions of second-order languages or the “inconvenient” semantic properties—incompleteness and noncompactness—of second-order logic. I suggest that the considerations of the present article preempt such reasoning. It can be agreed that, all things equal, it would be desirable to have a recursively axiomatized, compact logic with fewer presuppositions. I take it, however, that the purpose of logic is to study and codify correct inference. Since one cannot codify the correct inferences of a second-order language with a first-order logic, it follows that the logic of mathematics cannot be first-order.

Thus, I suggest that the presuppositions (and inconveniences) of second-order logic must be accepted. The purpose of this section is to briefly assess the presuppositions of second-order logic vis-à-vis first-order logic. The interested reader is referred to Boolos [1975] for more detailed considerations.

As noted by Boolos [1975], the present evaluation is simplified by the fact that in one important respect, the semantics of a given first-order language $L_1$ is the same as that of a corresponding second-order language $L_2$: a model or interpretation of either language consists of a nonempty domain together with a function giving appropriate assignments to the nonlogical constants. Since, in the cases at hand, the nonlogical terminologies are identical, the respective languages have exactly the same classes of interpretations. (Of course, this does not mean that a given theory in $L_2$ will have the same models as a counterpart in $L_1$.) The logics of $L_1$ and $L_2$ can therefore be evaluated in terms of the presuppositions of each concerning the common semantics.

4.1. Ontology. The main difference between the languages is that the second-order $L_2$ has variables ranging over the subsets of the domain, functions from the domain to the domain, etc. Thus, for a given interpretation, the first-order $L_1$ presupposes only the elements of the domain, while $L_2$ also presupposes the subsets, functions, and relations on the same domain.

I consider first the presuppositions of “pure” logic or, in other words, the presuppositions of uninterpreted languages. The logical truth $\exists x(x = x)$ of $L_1$
corresponds to a commitment to at least one element (of each domain). Similarly, the pair of logical truths $\exists X(\forall x(Xx))$ and $\exists X\forall x(\neg Xx)$ of $L2$ corresponds to a commitment to at least two subsets (of each domain), an empty set and a "universal" set. However, since there are interpretations whose domains have only a single element, the following is not a logical truth of $L2$:

$$\exists X \exists x \exists y (Xx \& Xy \& x \neq y).$$

That is, second-order logic uninterpreted does not presuppose a two-element set. Thus, the existential assumptions of uninterpreted second-order logic are rather weak and, moreover, are not much greater than those of uninterpreted first-order logic.

Of course, the concern with the ontological commitments of second-order languages is focused on the commitments of interpreted theories. For example, the concern with second-order arithmetic is directed at the commitment to sets of numbers. In general, let $d$ be the domain of an interpretation of both $L1$ and $L2$. Assume that $d$ has infinite cardinality $\kappa$. Concerning this interpretation, $L1$ is committed to the elements of $d$ or, in other words, to $\kappa$-many items. The further commitments of the second-order $L2$ are the subsets, etc., of $d$, which total $2^\kappa$ items. Thus, the second-order language as interpreted does have greater presuppositions, but the difference is (in effect) limited to a single "powerset" operation.31

4.2. Completeness and satisfaction. I turn now to the relative presuppositions of various metatheorems for first-order logic and second-order logic. For $L1$, the completeness theorem is that every consistent set of sentences is satisfiable. This theorem and its proof depend on the axiom of infinity (in the metalanguage). Indeed, if the semantics of $L1$ contained only finite domains, then the respective logic would not be complete. The reason for this is that there are consistent first-order sentences which are satisfiable only in infinite domains. An example of such a sentence is:

$$(\text{In}) \quad \forall x \exists y Rxy \& \forall x \forall y \forall z (Rxy \& Ryz \rightarrow Rxz) \& \forall x(\neg Rxz).$$

It is important to be clear as to what is, and what is not, presupposed here. First-order languages, by themselves, do not presuppose infinite domains. As above, a language itself hardly presupposes anything. However, first-order languages make it possible to presuppose infinite domains. That is, when one adopts or asserts a sentence like (In) (or accepts the metatheory needed to prove the completeness theorem), then, and only then, one is committed to an infinite domain.

The situation concerning second-order languages is similar, but less modest. It is easy to see that there are consistent second-order sentences which are satisfiable only in very large domains. For example, if $Z$ is the conjunction of the axioms of second-order $\text{ZFC}$, then $Z$ is satisfiable only in domains of inaccessible cardinality.

---

31 Although present concern is not with deductive systems, it might be noted that the soundness theorem for standard second-order logic does not require extensive set-theoretic presuppositions. Most of the axioms (and rules of inference) are either those of first-order logic or second-order versions of the quantifier axioms (which involve virtually no set theory). The only exceptions are the comprehension scheme and, perhaps, the axiom of choice. The soundness of these, of course, involves the axiom of separation and the axiom of choice, respectively, in the metalanguage.
I suggest that this is not a defect of second-order logic, but rather a result of the expressive ability of second-order languages. In other words, the present observations are the result of what I take to be the main advantage of second-order languages. As with the first-order case, the second-order $L_2$ does not, of itself, presuppose large domains, but does make it possible to presuppose large domains. The presuppositions come when one adopts or asserts a sentence like $Z$. In such a situation, and only then, one is committed to an inaccessible domain.

4.3. Logical truth. I close with an admission that there are statements of second-order logic that amount to rather substantial set-theoretic propositions. An example follows.

It is easily seen that there is a second-order formula $C(X)$ which (in any interpretation) is equivalent to “$X$ is either finite or denumerably infinite”, and that there is a second-order formula $E(X, Y)$ equivalent to “$X$ has the same cardinality as $Y$”. It follows that the formula

$$(O(X)) \quad C(X) \land \forall Y (Y \subseteq X \rightarrow C(Y) \lor E(X, Y))$$

is equivalent to “$X$ has cardinality $\aleph_1$”. There is also a second-order formula $P(X)$ equivalent to “$\exists x \exists y \exists f \exists g (\langle X, x, y, f, g \rangle$ is a model of real analysis)”. Of course, $P(X)$ amounts to “$X$ has the cardinality of the continuum”.

Consider the sentences

$$(CH) \quad \forall X (O(X) \equiv P(X)),$$

$$(NCH) \quad \forall X (P(X) \rightarrow \neg O(X)).$$

Notice that (CH) is a logical truth if and only if the continuum hypothesis is true, and that (NCH) is a logical truth if and only if the continuum hypothesis is false.

Of course, one does not normally think of the continuum hypothesis or its negation as a logical truth. The fact that one of them is a second-order logical truth is, again, a result of the expressive power of second-order languages: substantial statements about the semantics of $L_2$ can be made by sentences of $L_2$. Once again, this expressive power is here taken to be the main strength of second-order languages. The proper conclusion, I suggest, is not to reject second-order languages and second-order logic, but rather to reject the notion of a sharp distinction between mathematics and the logic of mathematics.

Acknowledgments. I would like to thank the following people for many helpful suggestions on earlier versions of this paper: John Corcoran, Haim Gaifman, Nicolas Goodman, Charles Kielkopf, George Kreisel, Timothy McCarthy, Jon Pearce, Michael Resnik, Michael Scanlon, George Schumm, Craig Smoryński, and an anonymous referee. I would also like to thank as a group the Ohio State philosophy of mathematics seminar for devoting several sessions to this project.

REFERENCES

J. Barwise [1977], *An introduction to first-order logic*, Handbook of mathematical logic (J. Barwise, editor), North-Holland, Amsterdam, pp. 5–46.

