Completeness and Categoricity: 19th Century Axiomatics to 21st Century Semantics

Steve Awodey and Erich H. Reck

July 10, 2001

Technical Report No. CMU-PHIL-118

Philosophy

Methodology

Logic

Carnegie Mellon

Pittsburgh, Pennsylvania 15213

Completeness and Categoricity: 19th Century Axiomatics to 21st Century Semantics

Steve Awodey
Department of Philosophy
Carnegie Mellon University
Pittsburgh, PA 15213
awodey@cmu.edu

Erich H. Reck
Department of Philosophy
University of California
Riverside, CA 92521
reck@citrus.ucr.edu

July 10, 2001

Abstract

This paper discusses several notions of completeness for systems of mathematical axioms, with special focus on their interrelations and historical origins in the development of the axiomatic method. We argue that higher-order logic is an appropriate framework for such systems, and we consider some open questions in higher-order axiomatics. In addition, we indicate how one can fruitfully extend the usual set-theoretic semantics so as to shed new light on the relevant strengths and limits of higher-order logic.

Contents

Introduction

One of the guiding tasks of 20th century mathematics and logic was that of *axiomatizing* mathematical concepts and even whole fields. This was part of the trend toward increasing systematization and abstraction in modern mathematics. Accordingly, the various possible notions of *completeness* of a system of axioms have taken on considerable interest, and their development in the late 19th and early 20th century now invites a historical review. This is true both for completeness understood as a property of logical calculi,[1] and for the quite different notion, or notions, of completeness as applying to axiomatic characterizations in mathematics generally, including the notion of categoricity.[2] Furthermore, recently several new technical results bearing on these issues have appeared. And finally, there even remain some open questions of a quite basic kind.

In this paper we address these issues systematically and comprehensively. In the first part of the paper, we document how the notion of categoricity and several related notions of completeness were first conceptualized. This occurred in connection with the development of the axiomatic method in late 19th and early 20th century mathematics, in the works of, among others, Richard Dedekind, Giuseppe Peano, David Hilbert, Edward Huntington, and Oswald Veblen. After the systematic development of formal logic there followed various logical and metamathematical investigations, exemplified by the well-known results of Kurt Gödel, Alfred Tarski, and others from the 1930s. Two further thinkers who contributed to these early metatheoretic investigations were Abraham Fraenkel and Rudolf Carnap, some of whose contributions actually predated those of Gödel and Tarski. Moreover, it was in Fraenkel's and Carnap's works from the 1920s that the most explicit, systematic comparisons of different notions of completeness can be found.[3]

A number of the questions formulated in the early metatheoretic works by Fraenkel, Carnap, and others still remain mathematically interesting today. It will become evident, however, that the now standard restriction to first-order logic in connection with them is both ahistorical and technically ill-advised. Topics like categoricity seem more naturally treated us-

[1]For recent discussions of the history of completeness as a property of logical calculi, see (Read 1997), (Sieg 1999), and (Zach 1999), earlier also (Goldfarb 1979), (Moore 1980), (Dreben and Heijenoort 1986), and (Moore 1988).

[2]Compare here (Corcoran 1980), (Corcoran 1981), and again (Read 1997). In the present paper we are, among others, answering some questions raised in (Corcoran 1981).

[3]The interesting role played by Carnap in this connection was established in (Awodey and Carus 2001). The present paper can be seen as a continuation of one topic discussed there.

ing higher-order logic, as was, in fact, done originally by Hilbert, Carnap, Gödel, Tarski, and others. In the second part of this paper, we give a concise introduction to this expanded logical framework, take up again the early metamathematical investigations, and pursue them in several directions. As a result, we provide partial answers to the questions mentioned earlier, and we indicate promising directions for further work.

Besides expanding the logical framework to that of higher-order logic, we also take a wider view of semantics than is customary, or was even possible until quite recently. Namely, we extend the range of semantic notions used from the standard set-theoretic semantics to more general topological and category-theoretic semantics. This might seem even more radical than the move to higher-order logic, but we believe it is justified by the light it sheds on some topics that have previously been obscure. It also allows us to establish some strengthenings of earlier results along lines hardly foreseeable by Carnap or Tarski, but not incompatible with their point of view.

Part I
Conceptual and Historical Background

1 Notions of completeness

Both for historical and logical purposes, it will be useful to start with an explicit distinction between several different notions of completeness. Assume in this connection that a formal language $\mathcal{L}$ is given, including the specification of the logical constructions allowed in the sentences of $\mathcal{L}$ e.g., propositional operations, quantification, higher types, etc. Assume also that notions of formal deduction and deductive consequence, on the one hand, and of interpretation, satisfaction, model, and semantic consequence, on the other, have been introduced in the usual way. This allows us to consider, in a mathematically precise way, whether a sentence φ is *deducible* from a set of sentences Γ (written $\Gamma \vdash \varphi$, also expressed by saying that Γ *yields* φ); whether some structure M *satisfies* a sentence φ (written $M \models \varphi$); whether M is a *model* of Γ (in the sense of satisfying all the sentences in Γ); and finally, whether Γ semantically *implies* φ (written $\Gamma \models \varphi$, and meaning that all models M of Γ satisfy φ).

Given such a syntax and semantics for $\mathcal{L}$ we can formulate the following definitions:

Definition 1. The deductive consequence relation $\vdash$ is called *complete* relative to the semantic consequence relation $\models$ if for all sentences φ and all sets of sentences Γ of $\mathcal{L}$: If $\Gamma \models \varphi$, then $\Gamma \vdash \varphi$.

Put informally, a deductive system is complete if it is "strong enough" for the corresponding semantics in the sense that it yields all the semantic consequences also as deductive consequences. As is well known, the standard deductive consequence relations for propositional and first-order logic are complete in this sense relative to conventional truth-value and set-theoretic semantics.[4] In contrast, no deductive consequence relation, in the usual sense, for second- or higher-order logic can be complete relative to standard set-theoretic semantics.[5]

Quite distinct but equally important are several notions of completeness for *mathematical theories* $\mathbb{T}$. Logicians today are accustomed to talking about a "theory" in three related senses: as a set of axioms (perhaps finite or recursively enumerable) formulated in terms of the primitive notions of some language $\mathcal{L}$ (the traditional mathematical notion of "axiomatic theory"); as the closure in $\mathcal{L}$ of a given set of sentences under either deductive or semantic consequence (the now-standard logical notion of "theory"); and as the set of all the sentences of $\mathcal{L}$ satisfied in some particular structure M (the "theory of M"). In the historical examples below, it is theories in the first of these senses—given by finitely many axioms—that are at issue. But the following definitions apply to all three kinds of theories:

Definition 2. A theory $\mathbb{T}$ is called *categorical* (relative to a given semantics) if for all models M, N of $\mathbb{T}$, there exists an isomorphism between M and N.

Informally, the idea here is that $\mathbb{T}$ has "essentially only one model". Familiar, examples are second-order Peano arithmetic, with the usual second-order induction axiom, and the second-order theory of a complete ordered field. In contrast, their usual first-order versions are not categorical.[6]

Two further familiar notions of completeness for a theory $\mathbb{T}$ are captured in the next two definitions. In them we will use the terms "semantically complete", "deductively complete", and "logically complete" in ways that are

[4]See the completeness theorems in (Bernays 1918), (Post 1921), (Gödel 1930), (Henkin 1950), and the historical discussions in (Sieg 1999) and (Zach 1999).

[5]By Gödel's incompleteness theorem; see (Gödel 1931). By "standard set-theoretic semantics" we mean to exclude Henkin models.

[6]For a discussion of categoricity in connection with examples related to Peano arithmetic, compare (Corcoran, Frank, and Maloney 1974).

not altogether standard. The reason for our choices will hopefully become evident.

Definition 3. A theory $\mathbb{T}$ is called *semantically complete* (relative to a given semantics) if any of the following equivalent conditions holds:

1. For all sentences φ and all models M, N of $\mathbb{T}$, if $M \models \varphi$ then $N \models \varphi$.
2. For all sentences φ, either $\mathbb{T} \models \varphi$ or $\mathbb{T} \models \neg\varphi$.
3. For all sentences φ, either $\mathbb{T} \models \varphi$ or $\mathbb{T} \cup \{\varphi\}$ is not satisfiable.
4. There is no sentence φ such that both $\mathbb{T} \cup \{\varphi\}$ and $\mathbb{T} \cup \{\neg\varphi\}$ are satisfiable.

Informally, the idea in 3.1 is that all models of the theory are "logically equivalent", in the sense that exactly the same sentences are satisfied by all of them (in the first-order case: elementary equivalent). The idea in 3.2 is that every sentence of the language is "semantically determined" by $\mathbb{T}$, so that either it or its negation is a semantic consequence of $\mathbb{T}$ (*tertium non datur*). Both second-order Peano arithmetic and the second-order theory of a complete ordered field are semantically complete, while their usual first-order versions are not. Tarski's theory of real arithmetic (the first-order theory for real closed fields) is semantically complete, but unlike the previous examples it is not categorical.[7]

Turning now to the deductive or syntactic side:

Definition 4. A theory $\mathbb{T}$ is called *deductively complete* (relative to a given deductive consequence relation $\vdash$) if any of the following equivalent conditions holds:

1. For all sentences φ, either $\mathbb{T} \vdash \varphi$ or $\mathbb{T} \vdash \neg\varphi$.
2. For all sentences φ, either $\mathbb{T} \vdash \varphi$ or $\mathbb{T} \cup \{\varphi\}$ is inconsistent.
3. There is no sentence φ such that both $\mathbb{T} \cup \{\varphi\}$ and $\mathbb{T} \cup \{\neg\varphi\}$ are consistent.[8]

Informally, the idea in 4.1 is that every sentence of $\mathcal{L}$ is "deductively determined" by $\mathbb{T}$, in the sense that either it or its negation is a deductive consequence of $\mathbb{T}$ (*tertium non datur*). Neither first- nor second-order Peano

[7] (Tarski 1951), compare also the discussion in (van den Dries 1988).

[8] By "consistent" and "inconsistent" we always mean deductively consistent and deductively inconsistent. Instead of "semantically consistent" we use "satisfiable" (as above).

arithmetic is deductively complete, likewise for the first- and second-order theories of a complete ordered field. On the other hand, Tarski's theory of real arithmetic provides an example that is not only semantically, but also deductively complete.

Clearly Definitions 4.1–4.3 are the deductive analogues of 3.2–3.4. It is also not hard to see that Definitions 4 and 3 are equivalent against the background of any logical system in which the deductive consequence relation is (sound and) complete in the sense of Definition 1, such as in the case of first-order logic. On the other hand, this is not true in general, as the second-order examples above illustrate. Note, in addition, that each of the notions introduced in Definitions 2, 3, and 4 is relative in a certain way: categoricity and semantic completeness to a corresponding semantics, and deductive completeness to a corresponding deductive system.

For historical purposes it will be useful to add two further, less familiar notions of completeness for a theory $\mathbb{T}$:

Definition 5. Let S be a set of sentences in the language $\mathcal{L}$ and let $\mathbb{T}$ be a theory in $\mathcal{L}$. $\mathbb{T}$ is called *relatively complete* (relative to S) if every sentence $\varphi \in S$ is provable from $\mathbb{T}$.

One can consider both informal and formal versions of this notion, relying on either an informal mathematical notion of proof or on provability as tied to some formal deductive system. We will later encounter several historical examples illustrating this notion. To anticipate, in them S will be the theorems of a certain field at a particular point in time, e.g. those of Euclidean geometry around 1900, and $\mathbb{T}$ will be a then-new set of axioms, such as Hilbert's.[9]

Finally, if we let $S = \{\varphi : \mathbb{T} \models \varphi\}$ in the previous definition:

Definition 6. A theory $\mathbb{T}$ is called *logically complete* (relative to a given semantics) if for all sentences φ, if $\mathbb{T} \models \varphi$ then φ is provable from $\mathbb{T}$.

One can evidently again consider both informal and formal versions of this notion, depending on whether one works with informal mathematical proofs or with proofs in a formal deductive system. Note that if we work with the latter, we are back to a case of completeness of the deductive consequence relation in the sense of Definition 1, namely where the parameter Γ is replaced by a particular theory $\mathbb{T}$. By way of example, even though higher-order deduction is not complete in the sense of Definition 1, it is not

[9]The historical importance of relative completeness, especially in connection with Hilbert, was pointed out to us by Wilfried Sieg. He calls it "quasi-empirical completeness".

hard to find a specific theory in higher-order logic that is logically complete in the sense of Definition 6, e.g., that of the notion of a set of some *particular* finite cardinality.

We consider next how these notions of completeness arose historically, namely in connection with the development of the axiomatic method in late nineteenth and early twentieth century mathematics.

2 Formal axiomatics

The use of the axiomatic method in mathematics goes back at least as far as Euclid's *Elements*, thus to around 300 BC. Traditionally, axiomatics was a method for organizing the concepts and propositions of an existent science in order to increase certainty in the propositions and clarity in the concepts. However, we are interested in a characteristically modern refinement of it, what is now often called *formal axiomatics*, earlier also *postulate theory*. In formal axiomatics the purpose is not primarily to increase certainly, nor is it merely to clarify and organize the concepts and theorems of a mathematical discipline in a systematic way. Rather, a further aim is to treat the objects of mathematical investigation more abstractly, and then to *characterize* them completely—to "define them implicitly", as it is often put somewhat misleadingly.[10]

Of course, the axiomatic method has also been applied very successfully in cases where such "completeness" of the axioms is not required, or even desirable, e.g. in the case of groups or topological spaces. In such cases it is not a matter of characterizing one particular mathematical structure but of studying various different, non-isomorphic, systems all satisfying certain general constraints. Thus in general, notions of completeness arise in contexts where axiomatizations are being undertaken with specific goals in mind. To say that an axiomatization is complete is, then, to say that the axiomatizers have achieved their goal, in particular that no further addition of "new axioms" is called for.

In its mature mathematical form, formal axiomatics involves using a *formal language*, a language that is taken to be uninterpreted and for which various different interpretations can be considered and compared. Ideally, at least in principle, formal axiomatics also requires making explicit which logical inferences between sentences of the language are permitted. This is usually done by specifying a *formal deductive system* that makes reference

[10] Compare (Corcoran 1995) for the goals of axiomatics. The name "formal axiomatics" as well as an influential endorsement of it go back to (Hilbert and Bernays 1934), p. 2.

only to the formal language and not its various interpretations.

We will now consider five historical examples of formal axiomatics which, in our view, represent the steps most relevant in its development. These examples are also closely linked to each other, as will become apparent.

2.1 Dedekind and Peano on the natural numbers

An important precursor, to some degree also a first example, of formal axiomatics in the sense just described is the treatment of the natural numbers and of elementary arithmetic in Richard Dedekind's "*Was sind und was sollen die Zahlen?*" from 1888.[11] In this classic essay Dedekind's goal is to put the theory of natural numbers on a new, uniform, and "logical" foundation. What that goal amounts to is explained in a well-known letter to the mathematician Keferstein, from 1890:

> What are the mutually independent fundamental properties of the sequence N, that is, those properties that are not derivable from one another but from which all others follow? And how should we divest these properties of their specifically arithmetic character so that they are subsumed under more general notions and under activities of the understanding *without* which no thinking is possible at all, but *with* which a foundation is provided for the reliability and completeness of proofs and for the construction of consistent notions and definitions?[12]

Note here Dedekind's emphasis on "completeness of proofs". This phrase reflects his goal to avoid any implicit, hidden assumptions in his proofs, thus to make explicit everything that is (and is not) relevant in the mathematical concepts involved. It also echoes the opening line of the Preface (first edition) to "*Was sind und was sollen die Zahlen?*", where Dedekind affirms: "In science nothing capable of proof ought to be accepted without proof."[13]

The "more general notions" Dedekind wants to use in giving a foundation to arithmetic are those of an informal theory of functions and sets; the latter he calls "systems". On their basis he proceeds to introduce various general conditions, or concepts, that such systems may satisfy. The central concept

[11](Dedekind 1888).

[12](Dedekind 1890), pp. 99–100. We are grateful to George Weaver for drawing this passage to our attention.

[13](Dedekind 1963), p. 31. In general, we use standard English translations of German texts in this paper, but occasionally we amend them.

is that of a "simply infinite system". In current terminology, its definition is this:[14]

Definition 7. A set S is said to be *simply infinite* if there exists a function f on S and an element $a \in S$ such that the following hold:

1. $f(S) \subseteq S$, i.e., f maps S into itself.
2. $a \notin f(S)$, i.e., a is not in the image of S under f.
3. $f(x) = f(y)$ implies $x = y$, i.e., f is a 1-1 function [Dedekind: f is *similar*].
4. S is the smallest set containing a and closed under f, i.e., it is the intersection of all such sets [Dedekind: S is *the chain under f with base point a*].

It is not hard to recognize what are now called the "Peano Axioms" (or "Dedekind-Peano Axioms") for the natural numbers in Dedekind's definition. A contemporary logical formulation—not much different from the original one in Giuseppe Peano's *Arithmetices Principia, Nova Methodo Exposita* of 1889[15]—is as follows: Taking N to be a set, s a function defined on N, and $1 \in N$,

1. $\forall x \; (x \in N \rightarrow s(x) \in N)$
2. $\forall x \; (x \in N \rightarrow 1 \neq s(x))$
3. $\forall x \; \forall y \; (s(x) = s(y) \rightarrow x = y)$
4. $\forall X \; [(1 \in X \wedge \forall y \; (y \in X \rightarrow s(y) \in X)) \rightarrow N \subseteq X]$

Note that this formulation uses second-order logic insofar as the induction axiom 4 uses a quantifier $\forall X$ over all sets. This corresponds to Dedekind's informal version which involves quantification over sets implicitly, but crucially in his clause 4.

Unlike Peano, Dedekind does not talk about "axioms" in his essay. Instead, he simply works with the concept of being a "simply infinite system" as defined above. He then introduces (as the result of a process of "abstraction") a particular simply infinite system N, with "base element" 1 and "set

[14] *Ibid.*, Definition 71. In the following passage we have not only amended the translation, but also updated the terminology and changed the order of Dedekind's four clauses.

[15] (Peano 1889), translated as (Peano 1973). Besides minor variations in the notation, Peano's version differs essentially only insofar as he includes axioms for equality as well.

in order" by ϕ, which he calls "the natural numbers"[16] After that, he proves a number of corresponding results, including the following two:

> *Theorem 132*: All simply infinite systems are similar [i.e., isomorphic] to the number series N and consequently [...] also to one another.
>
> *Theorem 133*: Every system which is similar to a simply infinite system and therefore [...] to the number series N is simply infinite.[17]

Dedekind does not yet work with a completely general notion of isomorphism, nor does he use the term "categorical". Nevertheless, these two theorems (and their proofs) show that he basically knows his characterization to be categorical. He then adds:

> *Remark 134*: [It is clear that] every theorem regarding numbers, i.e., regarding the elements n of the simply infinite system N set in order by the mapping ϕ, and indeed every theorem in which we leave entirely out of consideration the special character of the elements n and discuss only such notions as arise from the arrangement of ϕ, possesses perfectly general validity for every other simply infinite system S set in order by a mapping ψ and its elements s [...].[18]

The following related aspects of this remark are crucial for our purposes: First, Dedekind also realizes the semantic completeness of his axiomatization, essentially in the sense of our Definition 3.1 above. Moreover, he apparently infers this completeness directly from categoricity. At the same time, he presents these insights merely in the form of a "remark", not a "theorem", and he does not provide a proof. Indeed, giving such a proof would have required a more developed theory of logical syntax than he had at his disposal. Strictly speaking, Dedekind does not even work with the notion of a formal, uninterpreted language and corresponding interpretations

[16] *Ibid.*, Definition 73. For our purposes it does not matter how exactly Dedekind thinks about N, only that it is a particular simply infinite system. Compare (Tait 1997) for an interesting discussion of Dedekind's approach, especially of his notion of "abstraction".

[17] (Dedekind 1963), pp. 92 and 93, respectively.

[18] *Ibid.*, p. 95.

for it. Instead he talks about "translating" between the language for N and those for other simple infinities Ω.[19]

After having essentially established both categoricity and semantic completeness, in the rest of his essay Dedekind goes on to establish and provide the following: the general possibility of giving inductive definitions and proofs in arithmetic; specific inductive definitions for addition, multiplication, and exponentiation; proofs of the corresponding commutative, associative, and distributive laws; and a clarification of how to apply the natural numbers, as defined by him, to measure the cardinality of finite sets. What he establishes thereby, implicitly, is the relative completeness of his axioms in the sense of our Definition 5, here with respect to the usual, basic results in the arithmetic of natural numbers.

Finally, the overall structure of "*Was sind und was sollen die Zahlen?*" shows that Dedekind considers both the categoricity (derivatively also the semantic completeness) and the relative completeness of his characterization as conditions of adequacy for a systematic approach such as his. It is in these respects, or to that extent, that his work on the natural numbers should be counted as an early example of formal axiomatics. In other respects, however, his approach may be seen to be more "conceptual" than "formal", in particular insofar as he still lacks the notion of a formal language in the strict sense. And he is certainly a long way from a system of formal deduction that would allow the consideration of deductive completeness in the sense of our Definition 4.[20]

2.2 Hilbert on Euclidean space

Probably the most influential early example of formal axiomatics was David Hilbert's *Grundlagen der Geometrie*, first published in 1899.[21] In fact, it was this text that established the fruitfulness of such an approach in the mathematical community at large. *Grundlagen* starts as follows:

> Geometry, like arithmetic, requires only a few and simple princi-

[19] This last point is emphasized in (Corcoran 1981). At the same time, Dedekind clearly intends various different systems to fall under the concept "simply infinite system". He even considers systems that satisfy only some of the four clauses in it but not others; see, e.g., (Dedekind 1890), pp. 100–1.

[20] Frege's *Begriffsschrift* from 1879 could have provided some of the required notions and technical tools for Dedekind. But by his own account, Dedekind was unfamiliar with Frege's work at the time of writing "Was sind und was sollen die Zahlen?"; compare the prefaces to the first and second edition of (Dedekind 1963).

[21] We will use (Hilbert 1971), with corrections, but we will also have occasion to go back to (Hilbert 1903), (Hilbert 1902), and even (Hilbert 1899).

ples for its logical development. These principles are called the *axioms* of geometry.[22]

Of course, geometry had been axiomatized since the time of Euclid, as Hilbert immediately acknowledges. What is distinctive about his own approach is that it is self-consciously more abstract and "formal" than earlier ones. This does not mean that Hilbert has no intended interpretation or model for it in mind; in particular, he indicates that his choice of axioms is guided by a "logical analysis of our perception of space" (*ibid.*). What it means, instead, is that a central new method used by him is to consider a broad range of different interpretations, not only for his axiomatic system as a whole, but also for various parts of it (primarily to establish independence results). That is to say, Hilbert in effect treats the language of geometry as a formal language.[23] Along these lines, chapter one of *Grundlagen* starts with the following abstract description of its subject matter:

> *Definition:* Consider three distinct sets of objects. Let the objects of the first set be called *points* and denoted $A, B, C, \ldots$; let the objects of the second set be called *lines* and be denoted $a, b, c, \ldots$; let the objects of the third set be called *planes* and be denoted $\alpha, \beta, \gamma, \ldots$. [...] The points, lines, and planes are considered to have certain mutual relations, denoted by words like "lie", "between", "congruent". The precise and mathematically complete description of these relations follows from the *axioms of geometry*.[24]

Besides setting the stage for Hilbert's more "formal" approach, what is of greatest interest for us in the passage just quoted is his phrase "complete description". This phrase is, in fact, an echo of what Hilbert writes already in the Introduction of the work, where he states his goals as follows:

> This present investigation is a new attempt to establish for geometry a *complete*, and *as simple as possible*, set of axioms and

[22](Hilbert 1971), p.2.

[23]We say "in effect" because Hilbert still doesn't have an explicit, mathematically precise notion of interpretation *à la* Tarski at his disposal; moreover, compare the next footnote.

[24]*Ibid.*, p. 3, original emphasis. It is, we should note, still possible to read this definition as introducing an interpreted language, in such a way that it allows for various "reinterpretations", along the lines of Dedekind's "translations" of the language of the natural numbers. Hilbert will be considerably more definite about using formal languages in his later work.

> to deduce from them the most important geometric theorems in such a way that the meaning of the various groups of axioms, as well as the significance of the conclusions that can be drawn from the individual axioms, come to light.[25]

Throughout *Grundlagen*, Hilbert does not elaborate much on what he means by "completeness" in passages such as these. It is clear from the above, however, that he takes to be one of his primary goals what we have called relative completeness, namely with respect to "the most important geometric theorems" recognized by the mathematicians of his time.

To determine further what Hilbert could have meant by "completeness" in *Grundlagen*, we need to look more closely at his axioms and the roles they play in the work. These axioms are divided into five groups: (I) Axioms of Incidence, (II) Axioms of Order, (III) Axioms of Congruence, (IV) Axiom of Parallels, and (V) Axioms of Continuity. The two crucial ones for present purposes form group (V):

> V.1 (Archimedes' Axiom) If AB and CD are any segments, then there exists a number n such that n segments CD constructed successively from A on, along the ray from A through B, will pass beyond the point B.
>
> V.2 (Axiom of Line Completeness) It not possible to extend the system of points on a line with its order and congruence relations in such a way that the relations holding among the original elements as well as the fundamental properties of line order and congruence following from Axioms I-III and from V.1 are preserved.[26]

Later Hilbert add some explanations about the respective roles of these two axioms and about their relation to each other:

> The [line] completeness axiom is not a consequence of Archimedes' Axiom. In fact, in order to show with the aid of Axioms I–IV that this geometry is identical to the ordinary analytic "Cartesian" geometry Archimedes' Axiom by itself is not sufficient (cf. Sections 9 and 12). On the other hand, by invoking the [line] completeness axiom [...] it is possible to prove the existence of a limit that corresponds to a Dedekind cut as well

[25] *Ibid.*, p. 2, original emphasis.
[26] *Ibid.*, p. 26.

> as the Bolzano–Weierstrass theorem concerning the existence of condensation points; hence this geometry turns out to be identical to Cartesian geometry.[27]

And shortly thereafter:

> By the above treatment the requirement of continuity has been decomposed into two essentially different parts, namely into Archimedes' Axiom, whose role is to prepare the requirement of continuity, and the [line] completeness axiom which *forms the cornerstone of the entire system of axioms.* The subsequent investigations rest essentially only on Archimedes' Axiom and the completeness axiom is in general not assumed.[28]

Again later on in the text:

> [I]f in a geometry only the validity of the Archimedean Axiom is assumed, then it is possible to extend the set of points, lines, and planes by "*irrational*" elements so that in the resulting geometry on every line a point corresponds, without exception, to every set of three real numbers that satisfy its equation. By suitable interpretations it is possible to infer at the same time that *all* Axioms I–V are valid in the extended geometry. This extended geometry (by the adjunction of irrational elements) is thus none other than the ordinary space Cartesian geometry in which the [line] completeness axiom V.2 also holds.[29]

Several aspects in these remarks deserve comment: First, note that Hilbert is again explicit that his axioms allow for different interpretations or models. Thus, a "Cartesian" geometric space just based on the set of rational numbers and certain algebraic numbers fulfills all his axioms for Euclidean geometry besides the Axiom of Line Completeness.[30] Second, what that axiom adds is to insure that any system of objects satisfying all of the axioms is essentially the same as—in Hilbert's own words, "is none other than"—ordinary Cartesian space, as based on the set of real numbers. That

[27] *Ibid.*, p. 28.

[28] *Ibid.*, original emphasis.

[29] *Ibid.*, p. 59, original emphasis.

[30] As Hilbert points out, it suffices to consider the field of algebraic numbers that arise from the number 1 and the iterated application of five operations: addition, subtraction, multiplication, division, and the drawing of roots of the form $\sqrt{1+a^2}$ (*ibid.*, p. 29).

fact is presumably the sense in which for him it "forms the cornerstone of the entire system of axioms". In fact, what this last axiom does, against the background of the others, is to make Hilbert's whole system of axioms categorical.

At the same time, asserting simply and unequivocally that Hilbert understands his axioms to be categorical would be too strong. Note that, like Dedekind, he does not yet work with an explicit, general notion of isomorphism in *Grundlagen*. Moreover, he does not state a theorem that establishes, even implicitly, that his axioms are categorical; he leaves it at the short remarks above, without proofs. He also fails to observe that the semantic completeness of his axioms is a consequence. In the latter two respects his discussion actually falls behind Dedekind's. Finally, while relative completeness and (partial insights into) categoricity play some role in Hilbert's work, it never becomes entirely clear whether he means one or the other by the intended "completeness" of his system of axioms.

In fact, if we go slightly beyond *Grundlagen* it appears that what is meant by "completeness" in Hilbert's works from this period might be something else instead. In his article "*Über den Zahlbegriff*", published in 1900 and obviously written not long after *Grundlagen*, he comments again about the case of geometry:

> [In geometry] one begins by assuming the existence of all the elements [...] and then [...] brings these elements into relationship with one another by means of certain axioms [...]. The necessary task then arises of showing the *consistency* and the *completeness* of these axioms, i.e., it must be proved that the application of the given axioms can never lead to a contradiction and, further, that the system of axioms is adequate to prove all geometrical propositions. [...][31]

According to the last phrase in this passage, the axioms of geometry are supposed to allow for proofs of "all geometrical propositions", not just "the most important geometric theorems" as Hilbert wrote in *Grundlagen*. This opens up the possibility that what Hilbert really means by "completeness", both in "*Über den Zahlbegriff*" and in *Grundlagen*, is what we have called logical completeness: the (informal) provability of all truths of geometry from his axioms.

Overall it seems fair to say, however, that Hilbert is just not entirely clear on the notion of "completeness" at the time of writing *Grundlagen* and "*Über*

[31](Hilbert 1900), we use the translation (Hilbert 1996), pp. 1092–93, original emphasis.

den Zahlbegriff". Some passages in them perhaps point to categoricity (our Definition 2), others to relative completeness (Definition 5), and still others to logical completeness (Definition 6). In fact, the unclarity is furthered by Hilbert's use of the word "completeness" also in the "Axiom of Line Completeness", as well as by his practice of dropping the qualifier "line" in "line completeness" later on in the text.[32]

In connection with this additional use of "completeness" by Hilbert, two further clarifications should be made, one historical and one conceptual. First, the Axiom of Line Completeness is actually not present in the original German edition of *Grundlagen* from 1899. It can be found for the first time in the French translation of the text from 1900 (the year in which "*Über den Zahlbegriff*" appeared), after that also in the English translation from 1902, and then in the second German and all subsequent editions. Moreover, the initial version of the axiom is not that quoted above, but the following variant:

> *Axiom of Completeness.* It not possible to add new elements to a system of points, straight lines, and planes in such a way that the system thus generalized will form a new geometry obeying all of the five groups of axioms. In other words, the elements of geometry form a system which is incapable of being extended, provided that we regard the five groups of axioms as valid.[33]

That is to say, the Axiom of Completeness is initially formulated as a maximality condition for the whole space. It is only later that Hilbert reformulates it as a maximality condition just for lines in the space. (In later editions of *Grundlagen* the initial version of the Axiom of Completeness for the whole space becomes a theorem, i.e., is proved based on the axiom just for lines.[34])

The conceptual point of clarification is this: Hilbert's Axiom of Completeness asserts that (the whole space or) each line in space cannot be extended further—by adding additional points—while maintaining all of the other axioms. It is worth being very precise and explicit here so as to prevent a common misinterpretation. Namely, the axiom does *not* say anything about the semantic, deductive, or logical completeness of the system of axioms; *nor* does it say anything about categoricity, e.g., explicitly requiring

[32] The German word used in both cases is *Vollständigkeit*.

[33] (Hilbert 1902), p. 25.

[34] In a footnote Hilbert attributes this result to Paul Bernays; see (Hilbert 1971), p. 27.

the system of axioms to be categorical.[35] It is true, of course, that the Axiom of Line Completeness together with the other axioms has as a consequence the categoricity of the whole system of axioms; and such categoricity has, in turn, as a consequence the semantic completeness of this system of axioms. Still, what the Axiom of Line Completeness itself mentions is points in geometric space, not formulas in the corresponding language. In other words, what it asserts is the "completeness" (better: maximality) of the geometric space, not the completeness of the axiomatic system. This aspect comes out clearly if we reformulate Hilbert's axioms in formal logical terms. The Axiom of Line Completeness then shows itself to involve quantification over models of the axioms, not over sentences.[36]

Two final, related observations about *Grundlagen*: Like the Peano Axioms for the natural numbers, Hilbert's axioms for geometry can be formulated naturally and directly in higher-order logic. Indeed, except for Line Completeness, which is essentially higher-order, the axioms require only first-order logic. But Hilbert himself, like Dedekind before him, just works with an informal background theory of functions and sets. Second, at this point in time Hilbert, again like Dedekind, does not have a precise enough notion of formal deduction at his disposal to be able to conceptualize the notion of deductive completeness, as opposed to categoricity, semantic completeness, or informal logical completeness.[37]

2.3 Dedekind and Hilbert on the real numbers

Besides the natural numbers and geometric space, what called most urgently for an axiomatic treatment in 19th and early 20th century mathematics was the theory of the real numbers, and with it the Calculus. The contributions of three mathematicians are particularly interesting in this connection:

[35] It has been taken to do one or the other by various commentators, from (Veblen 1904), pp. 346-47 (see section 2.5 below), to (Zach 1999), p. 353. Compare Fraenkel (section 3.2 below) and (Corcoran 1972), p. 108, for clarifications concerning this issue.

[36] For interesting further discussions of Hilbert's Axiom of Line Completeness in the light of more general mathematical developments see (Ehrlich 1995) and (Ehrlich 1997); for more historical and philosophical background, in particular involving Hilbert's relation to Husserl in this connection, compare also (Majer 1997) and (DaSilva 2000).

[37] Strictly speaking, categoricity and semantic completeness involve the notion of semantic consequence, and that notion was also not given a fully explicit, mathematically precise articulation until the work of Tarski in the 1930s and 40s, perhaps even as late as the 50s; see (Hodges 1986). Nevertheless, an adequate informal understanding of the notion of semantic consequence can be seen to be implicit already in the writings by Dedekind and Hilbert considered so far, especially in connection with their treatment of independence questions.

Dedekind, Hilbert, and the American postulate theorist Edward V. Huntington.[38] We consider Dedekind's and Hilbert's contributions briefly in this section, and Huntington's in the next.

Today it is common to base the theory of the real numbers on the axioms for a complete ordered field. The first explicit version of these axioms can be found in Hilbert's "*Über den Zahlbegriff*" from 1900. However, considerations of the crucial component in them—a precise formulation of the axiom of line completeness or continuity—go back at least as far as Dedekind's "*Stetigkeit und Irrationale Zahlen*" from 1872.[39] What Hilbert did in "*Über den Zahlbegriff*" was not only to formulate his own version of that axiom, but to complement it with explicit axioms for an ordered field. Hilbert's axioms are divided into four groups, in analogy with his treatment of geometry: (I) Axioms of Composition (assuring the existence of sums, products, inverses, etc. for all numbers), (II) Axioms of Calculating (commutativity, associativity, etc.), (III) Axioms of Ordering (connecting addition and multiplication to the ordering, in the usual way); and finally, (IV) Axioms of Continuity.

Before examining the two axioms in Hilbert's group (IV), let us first remind ourselves of Dedekind's characterization of line completeness, as well as of some standard variants of it. Dedekind's main contribution in "*Stetigkeit und Irrationale Zahlen*" was to consider the following condition for a set of numbers R:

> *Dedekind continuity:* For all cuts (A, B) of R there is an element c in R such that $a \leq c \leq b$ for all $a \in A$ and all $b \in B$.

Given the axioms for an ordered field, this condition is equivalent to the following:

> *Least upper bound property:* For all subsets $S \subseteq R$, if S is bounded from above, then there is a least upper bound for S in R.

Several additional variants have also played an important role historically:

> *Bolzano continuity:* Every bounded, infinite subset of R has a condensation point in R.

[38] For general background on the "American postulate theorists", including Huntington, compare (Scanlan 1991).

[39] (Dedekind 1872).

> *Weierstrass continuity:* Every bounded, infinite, and increasing sequence of elements in R has a least upper bound in R.

> *Cauchy continuity:* Every infinite Cauchy sequence of elements in R converges to an element in R.

> *Cantor continuity:* Every infinite nested sequence of intervals in R has a non-empty intersection.

Each of these conditions captures, in a slightly different, but equivalent form, what it means for the real line to be "line-complete" or "continuous". An explicit logical formulation of any one of them requires second-order logic.

Hilbert, clearly aware of several of these alternatives, chooses none of them for his axiomatization of the reals. Instead, he uses the same procedure as in *Grundlagen*, taking as Axiom IV.1 the Archimedean Axiom and complementing it with the following:

> IV.2 (Axiom of Completeness): It is not possible to add to the system of numbers another system of things so that the axioms I, II, III, and IV.1 are also all satisfied in the combined system; in short the numbers form a system of things which is incapable of being extended while continuing to satisfy all the axioms.[40]

This condition might be abbreviated as follows:

> *Hilbert continuity:* There is no ordered Archimedean field of which R is a proper ordered subfield.

Hilbert is aware, again, that adding these two axioms rules out all unintended models. That is to say, he notes that any system of numbers satisfying all of his axioms is essentially the same as the familiar system of real numbers:

> Axioms IV.1 and IV.2 [...] imply (as one can show) Bolzano's theorem about the existence of a point of condensation. We therefore recognize the agreement of our number system with the usual system of real numbers.[41]

[40](Hilbert 1996), p. 1094
[41]*Ibid.*, p. 1095.

It is tempting, once more, to attribute a clear understanding of the categoricity of his axioms for the real numbers to Hilbert. However, as in the case of geometry there are reasons to be more circumspect and moderate in that respect. In particular, Hilbert does again not formulate a corresponding theorem, much less does he prove one; he only hints at the issue in the remark above. More basically, he still does not have a precise, general notion of isomorphism at his disposal. He also does not infer the semantic completeness of his axioms from the above.

Finally, Hilbert still has little to say about what he means by "completeness" in this case, except for the following brief, but pregnant remark at the very end of "*Über den Zahlbegriff*":

> Under the conception described above, the doubts which have been raised against the existence of the totality of all real numbers [...] lose all justification; for by the set of real numbers we do not have to imagine, say, the totality of all possible laws according to which the elements of a fundamental sequence can proceed, but rather—as just described—a system of things whose mutual relations are given by the *finite and closed* system of axioms I-IV, and about which new statements are valid only if one can derive them from the axioms by means of a finite number of logical inferences.[42]

Note the phrase "a finite number of logical inferences" at the end. This might be taken to point in the direction of formal deduction and the deductive consequence relation, although Hilbert still has no system of logical deduction at his disposal to give that notion a real bite. He also still shows no suspicion that there might be a difference between deductive and semantic consequence in general. Thus by "a finite number of logical inferences" he may also just mean an ordinary, informal mathematical proof.

2.4 Huntington on the positive real numbers

The next step in clarifying the notion of completeness, and in particular the relation between categoricity and semantic completeness, was taken by Edward V. Huntington in a series of papers from shortly after the turn of the century. The earliest and most relevant is his "A Complete Set of Postulates for the Theory of Absolute Continuous Magnitude" from 1902.[43]

[42] *Ibid.*, p. 1095.

[43] (Huntington 1902); compare also (Huntington 1903). Besides these two, a number of other articles on related topics (axioms for the complex numbers, groups, fields, etc.)

In this article, Huntington does not try to give axioms—or "postulates" as he prefers to call them—for the system of all real numbers, but only for the positive real numbers, which he calls "absolute continuous magnitudes". Besides relatively standard requirements for the algebraic and the ordering properties of the positive real numbers, this involves again an "axiom of continuity". Here is Huntington's version of it:

> *Postulate 5*: If S is any infinite sequence of elements a_k, such that $a_k < a_{k+1}$, $a_k < c$ $(k = 1, 2, 3, \ldots)$ (where c is some fixed element), then there is one and only one element A having the following two properties:
>
> 1. $a_k \leq A$ whenever a_k belongs to S;
> 2. if y and A' are such that $y + A' = A$, then there is at least one element of S, say a_r, for which $A' < a_r$.[44]

He adds in a footnote: "This postulate 5 is essentially the same as the principle employed by Weierstrass, in his lectures, for the definition of an irrational number." Thus Huntington does not use a Hilbert-style maximality condition, although he draws on Hilbert's work in some other ways.[45]

Early on in his essay Huntington writes about his goals:

> *Introduction:* The following paper presents a complete set of postulates or primitive propositions from which the mathematical theory of absolute continuous magnitude can be deduced. [...] The object [...] is to show that [the following six postulates] form a complete set; that is, they are (I) *consistent*, (II) *sufficient*, (III) *independent* (or *irreducible*). By these three terms we mean: (I) there is at least one assemblage in which the chosen rule of combination satisfies all the six requirements; (II) there is essentially *only one* such assemblage possible; (III) none of the six postulates is a consequence of the other five. [...] *[T]he propositions 1–6 form a complete logical basis for a deductive mathematical theory.*[46]

were published by Huntington in the *Transactions of the American Mathematical Society* during the following years. For summaries of the corresponding results see (Huntington 1911) and (Huntington 1917).

[44](Huntington 1902), p. 267. Both here and below we have changed Huntington's notation slightly.

[45]Huntington's paper contains a generous, interesting list of historical references, including to (Hilbert 1900); see (Huntington 1902), pp. 265-66.

[46]*Ibid.*, p. 266, original emphasis.

The second of Huntington's three conditions of adequacy for an axiomatic system—what he calls "sufficiency"—is clearly the one most relevant to the current discussion.

Like Dedekind in the case of the natural numbers, Huntington devotes several lemmas and theorems to his condition of "sufficiency" in the rest of his paper. The most important of them is the following:

> *Theorem II:* Any two assemblages M and M' which satisfy the postulates 1–6 are equivalent; that is they can be brought into one-to-one correspondence in such a way that $a + b$ will correspond with $a' + b'$ whenever a and b in M correspond with a' and b' in M' respectively.[47]

That is to say, what Huntington provides is this: a careful formulation of the notion of isomorphism; an explicit definition of categoricity ("sufficiency") based on it; and a separate theorem, with proof, to the effect that his system of postulates is categorical.

At the same time, what Huntington means by "completeness" in the passage from his Introduction above still remains somewhat unclear. Much depends on what is meant by his cryptic phrase "a complete logical basis for a deductive mathematical theory". There is no question that he makes categoricity central to his paper, which suggests that that is what he means by "completeness". However, the phrase "deductive mathematical theory" points to either deductive or logical completeness (in a formal or informal sense). Furthermore, any awareness that these notions might be significantly different from categoricity or semantic completeness is still missing in the paper.

Nevertheless, Huntington combines, in an explicit and careful way, several of Dedekind's and Hilbert's insights. He also coins—apparently for the first time—a special name for categoricity, namely "sufficiency". In those respects formal axiomatics is consolidated on a high level in his work, at least with respect to its semantic side.[48]

[47] *Ibid.*, p. 277.

[48] As John Corcoran has pointed out to us, a particularly interesting, systematic treatment of these issues can be found in (Huntington 1917), parts of which were published already in 1905–6. It would be worth analyzing Huntington's contributions in this and related works further.

2.5 Veblen on Euclidean and projective geometry

Hilbert's axiomatic approach, especially as applied to geometry, was also adopted and developed further by Oswald Veblen, another of the so-called American postulate theorists.[49] Veblen started his mathematical career with a detailed study of Hilbert's *Grundlagen*. As a result he proposed a modified set of axioms, first published in his "A system of axioms for geometry" from 1904.[50]

Several of the notions discussed so far come up in Veblen's paper. To begin with, in describing his goals he writes:

> It is part of our purpose to show that there is *essentially only one* class of which the twelve axioms are valid. [...] Consequently any proposition which can be made in terms of points and order either is in contradiction with our axioms or is equally true of all classes that verify the axioms. The validity of any possible statement in these terms is therefore completely determined by the axioms; and so any further axioms would have to be considered redundant. [...] A system of axioms such as we have described is called *categorical*, whereas one to which it is possible to add independent axioms (and which therefore leaves more than one possibility open) is called *disjunctive*.[51]

Regarding his terms "categorical" and "disjunctive" Veblen adds in a footnote: "These terms were suggested by Professor John Dewey." In the main text he continues:

> The categorical property of a system of propositions is referred to by Hilbert in his "Axiom der Vollständigkeit", which is translated by Townsend [the translator of *Grundlagen*] into "Axiom of Completeness". E.V. Huntington, in his article on the postulates of the real number system, expresses this conception by saying that his postulates are sufficient for the *complete definition* of essentially a single assemblage. It would probably be better to

[49]For more on Veblen's contributions to logic and the foundations of mathematics see (Aspray 1991); compare also again (Scanlan 1991).

[50](Veblen 1904), compare also (Veblen 1902). What Veblen tried to do, in particular, was to reduce the number of primitive notions in geometry to two: "point" and "order". However, this reduction didn't quite work, as one needs "congruent" in addition. Compare (Tarski and Lindenbaum 1926) and (Tarski 1983), pp. 306-07, for later clarifications about this issue. We are grateful to Michael Scanlan for clarifying this detail for us.

[51](Veblen 1904), p. 346, original emphasis.

> reserve the word *definition* for the substitution of one symbol for another, and to say that a system of axioms is categorical if it is sufficient for the complete *determination* of a class of objects or elements.[52]

A number of points in these two passages deserve our attention.

First, Veblen is obviously quite clear about what categoricity amounts to, referring back to Huntington in that connection. At the same time, when he writes that "the categorical property of a system of *propositions* is referred to by Hilbert in his '*Axiom der Vollständigkeit*' " he apparently misinterprets Hilbert's axiom, or at least describes it in a misleading way.

Second, Veblen, like Dedekind before him and again without proof, remarks explicitly that semantic completeness is a direct consequence of categoricity. Yet, note that his main formulation of semantic completeness—"any proposition either is in contradiction with our axioms or is equally true of all classes that verify the axioms"—does not amount to our Definition 3.1, as in Dedekind's case, but to Definition 3.3. (We are interpreting "in contradiction with our axioms" here as "not satisfiable together", i.e., as involving semantic inconsistency. We will provide further justification for that interpretations shortly.) In addition, Veblen's subsequent remark that "the validity of any possible statement in these terms is therefore completely determined by the axioms" agrees with Definition 3.2. And Veblen's definition of a system of axioms being "disjunctive"—"one to which it is possible to add independent axioms (and which therefore leaves more than one possibility open)"—points to Definition 3.4. So three of our four versions of "semantic completeness" come up explicitly in Veblen's remarks, and he treats them as obviously equivalent.

Third and perhaps most interestingly, Veblen relates categoricity more closely to semantic completeness than has been done previously. Note, e.g., how he introduces being disjunctive as a sort of complementary concept to—the negation of?—that of being categorical, and also as the negation of semantic completeness in the form of Definition 3.4. Still, it remains unclear what exactly the relation between these concepts is supposed to be.

In 1906 Veblen published another article on the same general topic, called "The foundations of geometry: a historical sketch and a simple example". This article was written for the magazine *Popular Science Monthly*, as an overview article for a broader audience. It contains several passages which illuminate Veblen's views further. In connection with the notion of categoricity he now remarks:

[52] *Ibid.*, pp. 346–47.

> If we have before us a categorical system of axioms, every proposition which can be stated in terms of our fundamental (undefined) symbols either is or is not true of the system of objects satisfying the axioms. In this sense it either is a consequence of the axioms or is contradictory with them.[53]

Let us suppose that what Veblen meant here was that "every proposition either is or is not true of *every* system of objects satisfying the axioms" (since, as he had emphasized earlier, a categorical system of axioms has "essentially only one" model). Then we can see him again moving without hesitation from categoricity to semantic completeness in this passage, the latter now formulated in the form of Definition 3.3—assuming we take the phrases "consequence of the axioms" and "contradictory" in the semantic sense.

That Veblen usually does mean "consequence" in the semantic sense in the articles under discussion is confirmed by another brief remark from his 1904 article. There he notes that in the case of a categorical, thus semantically complete, system "[any new axiom is redundant] even were it not deducible from the axioms by a finite number of syllogisms".[54] Note, at this point, the following: what Veblen suggests here is that a potential new axiom might be a semantic consequence of the old axioms without being a deductive consequence of them, i.e., without being "deducible in a finite number of syllogisms". What that implies, of course, is that the notion of semantic consequence might not coincide with that of deductive consequence. This is a radically new suggestion.

In another brief aside from his 1906 article Veblen is more direct and explicit, even if still somewhat hesitant, on the same topic. Here he formulates the following question:

> But if [a proposition] is a consequence of the axioms, can it be derived from them by a syllogistic process? Perhaps not.[55]

Given that Veblen, like Dedekind, Hilbert, and Huntington before him, is not using a precise notion of syntactic consequence, and only an implicit notion of semantic consequence, this question is quite remarkable and insightful. With it Veblen takes a significant step beyond all the other authors considered so far.

[53](Veblen 1906), p. 28.
[54](Veblen 1904), p. 346.
[55](Veblen 1906), p. 28.

A final word on Veblen: Soon after finishing his work on Hilbert and Euclidean geometry, he turned his attention to projective geometry; and within a few years he and his co-worker J.W. Young succeeded in, among other things, formulating a categorical system of axioms for that geometry as well.[56]

3 Logic and metatheory

Let us take stock briefly. By 1908 we have axiomatizations for several main areas of then-contemporary mathematics: the theories of the natural numbers, the real numbers, and Euclidean and projective geometry. In each case "completeness" is stated as an explicit goal, a criterion of adequacy for the axiomatization. What "completeness" means, more or less explicitly, is primarily categoricity, secondarily semantic completeness (in various equivalent forms), and in some cases even relative completeness or logical completeness. Also, semantic completeness is repeatedly recognized to be a direct consequence of categoricity, although no proof of that fact is ever given; and sometimes the two notions are conflated, or apparently treated as equivalent. Finally, it is only around 1904–6 that we have found the first expression of a suspicion, in some asides of Veblen's, that neither categoricity nor semantic completeness may need to coincide with deductive or logical completeness, or more generally that the deductive consequence relation may differ from its semantic counterpart.

3.1 *Principia Mathematica* and its descendants

From a contemporary point of view the main ingredient missing in the works considered so far is a precise and purely formal notion of deductive consequence. Without such a notion, it is hard to study the relation between semantic and deduction consequence systematically, or even to formulate the relevant questions in a precise and fruitful way. That situation only changed gradually. Ignoring the work of Gottlob Frege, as was in effect done at the time,[57] the first major step forward in that connection was the publication

[56] Velben and Young's axiom system for projective geometry was first published in (Veblen and Young 1908) and discussed more systematically in (Veblen and Young 1910).

[57] Frege's work on logic in (Frege 1879) and later writings failed to have a significant influence on the developments discussed so far, as already mentioned in the case of Dedekind. This was, no doubt, partly due to his traditional, anti-formalist views about axiomatics. For illuminating recent discussions of that aspect of Frege's logic see (Blanchette 1996) and (Goldfarb 2001).

of Whitehead and Russell's *Principia Mathematica* in 1910–13.[58] Although the authors of *Principia* did not cast their logic into a formal axiomatic mold in the spirit of Dedekind, Peano, Hilbert, Huntington, and Veblen, they did convince several mathematicians and logicians of the value of their new, more formal approach to logical deduction, notably Hilbert and Rudolf Carnap.[59]

The logic presented in *Principia* was essentially higher-order predicate logic, together with a controversial "ramified" theory of types and axioms of reducibility, infinity, and choice. From a later point of view, it contains a number of philosophically-motivated complications that were mathematically inconvenient and unnecessary. This was recognized gradually in the 1920s, in connection with the following two discoveries: First, one can isolate the subsystems of propositional and first-order logic and study them with good results. Second, one can simplify the higher-order part of the logic to the "simple" theory of types, thus also eliminating the need for the problematic axiom of reducibility, at least for mathematical purposes.

From today's point of view it hardly seems necessary to motivate the separate attention given to propositional and first-order logic. We have come to understand that these subsystems have interesting and mathematically significant properties. In particular, both propositional and first-order logic are complete with respect to standard truth-value and set-theoretic semantics, in the sense of Definition 1 above. For propositional logic this result was established independently by Paul Bernays, in an unpublished work from 1918, and by Emil Post, who published it in 1921.[60] For first-order logic it was established by Kurt Gödel in 1929.[61] Moreover, first-order logic was early-on shown to have various related characteristics like compactness and the Löwenheim-Skolem properties.

From the 1910s to the 1930s, most logicians working on axiomatics and the foundations of mathematics—including Hilbert, Gödel, Carnap, and Tarski—did not work with first-order logic, however, but with some version of higher-order logic, along the lines of simple type theory. A main historical source for that theory was Frank Ramsey's article "Mathemati-

[58] See especially the first volume, (Whitehead and Russell 1910).

[59] For Hilbert see (Sieg 1999); for Carnap see section 3.2 below.

[60] See (Bernays 1918) and (Post 1921); compare also the historical discussion in (Sieg 1999) and (Zach 1999).

[61] See (Gödel 1929) and the published version in (Gödel 1930); compare also (Henkin 1950). Around the same time as Gödel, and independently, Jacques Herbrand developed similar ideas in his dissertation; compare the historical notes in (Church 1956), p. 291, (Goldfarb 1971), p. 265ff., and (Dreben and Heijenoort 1986).

cal Logic" from 1926, in which various arguments for simplifying the logic of *Principia* were given.[62] Similar suggestions were also made by others around that time, including the Polish logician Leon Chwistek.[63] The first general expositions of the theory were published in Hilbert and Ackermann's *Grundzüge der Theoretischen Logik* from 1928 and, independently, in Rudolf Carnap's *Abriß der Logistik* from 1929. The theory reached its "canonical" form in Alonzo Church's "A formulation of the simple theory of types" from 1940.[64]

Of course, we now know that neither higher-order logic nor the restricted fragment called second-order logic are complete in the sense of Definition 1 with respect to their standard set-theoretic semantics, as was famously established by Gödel in 1930.[65] It should be kept in mind, however, that this incompleteness is relative to a particular choice of semantics.[66] Moreover, owing to its greater expressive capacity higher-order logic has some important advantages for axiomatics. In particular, it permits the finite and categorical axiomatization of the classical mathematical theories discussed above.

3.2 Fraenkel, Carnap, and early metatheory

In addition to the emergence of both first-order logic and the simple theory of types, the 1920s and 30s saw an increase of attention to metatheoretic questions, now also including consideration of formal deduction, and especially in connection with the notions of completeness and categoricity.

Much of this work came out of, or was influenced by, the Hilbert school of proof theory centered at Göttingen.[67] At this point, Hilbert and his

[62](Ramsey 1926). Ramsey's views were influenced by, among others, Ludwig Wittgenstein.

[63]See (Chwistek 1925), also (Chwistek 1967), especially pp. 342-43; the latter was originally published as (Chwistek 1921). Compare also again the corresponding notes in (Church 1956), p. 355.

[64]See (Hilbert and Ackermann 1928), (Carnap 1929), and (Church 1940). Note that (Gödel 1931) is also based on a version of the simple theory of types. Moreover, it should be emphasized that Frege's *Begriffsschrift* of 1879 already contained the essentials of simple type theory.

[65]The result was first published in (Gödel 1931). For historical notes in this connection see (Dreben and Heijenoort 1986).

[66]In Part II of this paper we will indicate an alternate semantics relative to which deductive higher-order logic is complete.

[67]See again (Sieg 1999). We also count Hermann Weyl as a member of Hilbert's school here; compare in this connection (Weyl 1926), chapter I, called "Mathematical Logic. Axiomatics".

coworkers payed special attention to deductive issues, thus going far beyond Hilbert's *Grundlagen der Geometrie* and "*Über den Zahlbegriff*" in that respect. As a result, influential statements of the question whether first-order logic is complete in the sense of our Definition 1 were published both in Hilbert and Ackermann's *Grundzüge der Theoretischen Logik* from 1928 and in Hilbert's "*Probleme der Grundlegung der Mathematik*" from 1929; and the same is true for the question whether the then usual axiom systems for the natural and real numbers are deductively or logically complete in the sense of our Definitions 3 and 6, respectively.[68] Answers to these questions, as well as further results along similar lines, were primarily due to Kurt Gödel in Vienna and to Alfred Tarski and his coworkers in Warsaw.[69]

Here we will focus on the contributions of two other figures: Abraham Fraenkel and Rudolf Carnap. Their works are particularly relevant for several reasons: First, many of their metatheoretic investigations actually predate those of Gödel and Tarski, and are largely independent of the Hilbert school. Second, there is a direct connection between their investigations and the developments described earlier in this paper. Third, unlike most metatheoretic studies from the 1930s and 40s on, theirs are not restricted to first-order logic, thus providing us with a useful broader perspective. And fourth, some of the questions raised in their writings—especially concerning the relation between semantic completeness and categoricity in the specific context of higher-order logic—are not only interesting, but also still unresolved. Overall, we believe that Fraenkel and Carnap deserve more attention and credit in this connection than they have received so far.

Probably the first text to focus directly and systematically on the relation between categoricity and several different notions of completeness was Fraenkel's *Einleitung in die Mengenlehre*. This book was initially published in 1919, enlarged to a second edition in 1923, and enlarged again to a third edition in 1928.[70] The first edition is still silent on this issue, but in the second edition, Fraenkel adds a separate section on "the axiomatic method". In it he considers several general questions and conditions concerning axiomatic theories of the kind we encountered above, i.e., finite sets of axioms. Thus he writes:

[68] (Hilbert and Ackermann 1928) and (Hilbert 1929); the latter was presented as a lecture in Bologna in 1928. For historical background compare here (Dreben and Heijenoort 1986) and (Mancosu 1998), pp. 149-88.

[69] For Gödel see the works cited above; for Tarski see many of the articles in (Tarski 1983), especially (Lindenbaum and Tarski 1935).

[70] For the second and third editions see (Fraenkel 1923) and (Fraenkel 1928). Translations of passages from these works will be our own.

Besides independence [of the axioms] a second, even more important property usually required, if possible, from a system of axioms is the *completeness of the system.* This property has been studied much less so far and, when studied at all, has not always been understood in the same sense. What probably comes to mind first is the conception according to which the completeness of an axiomatic system demands that the axioms encompass and govern the entire theory based on them, in such a way that every relevant question can be answered, one way or the other, by means of inferences from the axioms. Obviously assessing completeness in this sense is closely connected with the problem of the decidability of mathematical questions discussed in the previous paragraph (p. 169f.) [...] and is, thus, impeded by considerable difficulties. [...]

More sharply circumscribed and easier to assess is another sense of completeness for a system of axioms, a sense first characterized fully by *O. Veblen,* it seems.[71] According to it an axiomatic system is called complete if it determines uniquely the mathematical objects governed by it, including the basic relations between them, in such a way that between any two interpretations of the basic concepts and relations one can effect a transition by means of a 1-1 and isomorphic correlation. [...][72]

Thus in 1923 Fraenkel distinguishes clearly between categoricity (the second notion mentioned) and what looks very much like deductive completeness (the first notion mentioned). However, no distinction is made between deductive and semantic completeness, leaving a small doubt about what is meant by the phrase "inferences from the axioms" above.

Fraenkel adds an explicit discussion of the latter distinction in the third edition of his book. There the passage just quoted is modified and expanded as follows:

[T]he completeness of a system of axioms demands that the axioms encompass and govern the entire theory based on them in such a way that every question that belongs to and can be formulated in terms of the basic notions of the theory can be answered, one way or the other, in terms of deductive inferences from the

[71] In a footnote Fraenkel refers to (Veblen 1904) and (Huntington 1902) at this point, as well as to earlier work of his own.

[72] (Fraenkel 1923), pp. 226-27, original emphasis.

axioms. Having this property would mean that one couldn't add any new axiom to the given system (without adding to the basic notions) so that the system was "complete" in that sense; since every relevant proposition that was not in contradiction with the system of axioms would already be a consequence and, thus, not independent, i.e., not an "axiom". [...]

> Closely related to this first sense of completeness, but by far not as far reaching and easier to assess, is the following idea: [...] In general, a number of propositions that are inconsistent with each other and that can, thus, not be provable consequences of the same system of axioms can nevertheless be compatible with that system individually. Such a system of axioms leaves open whether certain relevant questions are to be answered positively or negatively; and it does so not just in the sense of deducibility by current or future mathematical means, but in an absolute sense (representable by independence proofs). A system of axioms of that kind is then, with good reason, to be called incomplete. As a consequence, one can [...] pose the problem of completeness also as follows: Let A be a proposition relevant with respect to a given system of axioms. The system is to be called *complete* if, no matter whether we in fact succeed to deduce the truth or falsity of A from the system or are able to secure its deducibility theoretically, only *either* the truth *or* the falsity of A—but not both possibilities—is *compatible* with the system. [...]
>
> Quite different, finally, is another sense of completeness, one probably characterized explicitly for the first time by Veblen.[73] [...] According to it a system of axioms is to be called *complete*—also "categorical" (Veblen) or "monomorph" (Feigl-Carnap)—if it *determines the mathematical objects falling under it uniquely in the formal sense*; i.e., such that between any two realizations one can always effect a transition by means of a 1-1 and isomorphic correlation.[74]

Clearly at this point, in 1928, Fraenkel is able to characterize distinctly first deductive completeness, then semantic completeness, and finally categoricity, along lines quite close to our Definitions 4, 3, and 2, respectively. Also,

[73] Here Fraenkel refers again, now in the text, not just in a footnote, to (Veblen 1904) and (Huntington 1902).

[74] (Fraenkel 1928), pp. 347–49, original emphasis.

with respect to both deductive and semantic completeness he mentions several of the variants distinguished by us and, like Veblen, recognizes their equivalence.

An further step forward in the 1928 edition of Fraenkel's book is his recognition and clarification of the difference between completeness in any of his three senses, on the one hand, and completeness in the sense of Hilbert's "Axiom of Completeness", on the other. Thus in a footnote, attached to the second paragraph quoted above, Fraenkel writes:

> So as to avoid misunderstandings let me emphasize that this kind of completeness [deductive completeness] has conceptually nothing to do with that involved in [Hilbert's] "Axiom of Completeness" [...]. In the latter it is the objects governed by the axioms, in the former the axioms themselves, that are not capable of extension. Of course, there is still a close connection between what is expressed in the Axiom of Completeness and the notions of completeness to be discussed below. This connection awaits clarification in detail. [...][75]

Fraenkel is obviously more careful and precise here than Veblen was several years earlier.

In the main text, Fraenkel continues with a further clarification of the relation between deductive and semantic completeness:

> If one compares the three different (and, incidentally, by no means exhaustive) notions of completeness above, completeness in the first sense has obviously a special status; it has, correspondingly, also been called "*Entscheidungsdefinitheit*". We could assess it only by "the establishment of a fixed method of proof that leads, provably, to the solution of any relevant problem" As such it is to be left aside as unrealizable if the area in question is not trivial, e.g., of strictly finite structure (Weyl [7], p. 20).[76] The situation is quite different with respect to the second notion. In that case there is, as we should note, a difference between a decision "being-determinate-in-itself" and the general *establishment* of what that decision is, e.g., in the form of a method of proof. Put in a more mathematical way: A system of axioms could actually determine an area insofar as never to

[75] *Ibid.*, p. 347.

[76] The reference is to (Weyl 1926).

> allow that besides a well known axiom A its contradictory opposite $\neg A$ is also compatible with the axioms, while at the same time a decision was impossible about whether A or $\neg A$ holds, e.g., because such a decision could not be forced in a finite number of steps! Moreover, the establishment of a general method to make such decisions could be impossible. In many cases the [semantic] completeness of a system of axioms may, then, be a fact. But the question of how to establish that fact—as a characteristic property of a system of axioms—is still open. That question is obviously of considerable interest, as is the question of how to connect it to completeness in the third sense above [categoricity].[77]

Two aspects of this last passage are particularly noteworthy: First, Fraenkel is much more clear and definite than Veblen—not to mention Dedekind, Hilbert, and Huntington—about the difference between deductive and semantic completeness. He is also strikingly pessimistic about the possibility of having a "non-trivial" system of axioms that is deductively complete (partly because, following Weyl, he still thinks it is not possible to come up with a logical calculus that is complete in the sense of our Definition 1). Second, at the end of the passage he explicitly poses the question of how semantic completeness and categoricity are related (in conjunction with the question of how to establish that a system is semantically complete in the first place). As we saw, several earlier writers had stated, without proof, that categoricity implies semantic completeness; but crucially, Fraenkel's question also concerns the converse: Is it the case that semantic completeness implies categoricity?

This is the point at which to turn to Rudolf Carnap, in particular to a neglected work on logic and axiomatics from the second half of the 1920s entitled *Untersuchungen zur Allgemeinen Axiomatik*.[78] In it Carnap extends Fraenkel's considerations in the following three ways: He makes serious attempts to answer Fraenkel's questions about the precise connections between categoricity, deductive completeness, and semantic completeness. Unlike Fraenkel, he puts the corresponding investigations into a formal, logical framework, namely that of the simple theory of types. And he picks up on Fraenkel's question concerning the relation between his three notions of

[77] *Ibid.*, p. 352, original emphasis.

[78] This work has only recently been edited and published, based on manuscripts found in Carnap's *Nachlaß*; see (Carnap 2000). In what follows we draw heavily on the study of it (Awodey and Carus 2001).

completeness, on the one hand, and completeness in the sense of Hilbert's "Axiom of Completeness", on the other.

Before considering Carnap's investigations further, some basic ideas and results need to be clarified from the point of view of a contemporary reader so as to prevent some possible confusions. To begin with, it is well-known today, and not hard to prove given the proper setup, that the categoricity of an axiomatic theory implies its semantic completeness. This is not only true in the case of first-order logic, but also for axiomatic theories in higher-order logic.[79] On the other hand, the question of whether the converse holds has not been answered completely even today, in spite of the fact that it is, to use Fraenkel's words, "obviously of considerable interest". In addition, this inference, from semantic completeness to categoricity, depends crucially on two background conditions: First, it depends on the logical language used, in particular on what sorts of sentences φ are supposed to occur in the definition of semantic completeness. Clearly the inference fails, e.g., if we restrict attention to just first-order sentences.[80] But what about the case of higher-order logic? Here, secondly, it is crucial to be precise about what is meant by "axiomatic theory". Indeed, it is not hard to see that the inference from semantic completeness to categoricity fails again if we consider general "theories" in the sense of arbitrary sets of sentences in some given language (by an argument from the bounded cardinality of such sets) . However, in the historical examples above we were concerned with the specific case of *finite* sets of axioms. The remaining question—arguably the one Fraenkel had in mind—is then this: For a theory $\mathbb{T}$ with finitely many axioms in higher-order logic, does the semantic completeness of $\mathbb{T}$ (in the sense of Definition 3 above) imply its categoricity (in the sense of Definition 2 above)?[81]

Answering this and some related questions was exactly the task that Carnap—who had not only studied the 1923 edition of Fraenkel's book carefully, but also contributed to its 1928 edition[82]—set himself during

[79] See (Lindenbaum and Tarski 1935), p. 390, for an early statement of this result; compare also section 4.4 below.

[80] As the Löwenheim-Skolem theorems imply, a first-order theory may have only one elementary equivalence class of models and yet not be categorical.

[81] Cutting to the chase, the answer to this question is still unknown. We will consider a few special cases for which we know the answer to be positive in section 4.4; compare also (Lindenbaum and Tarski 1935) in this connection.

[82] Carnap communicated his own research to Fraenkel between the second and third edition of Fraenkel's book, including Part I of (Carnap 1928). Besides Fraenkel's reference to Carnap's (and Feigl's) notion of "*Monomorphie*", see here the preface to (Fraenkel 1928) in which he thanks Carnap for his help, refers to (Carnap 1927), and mentions "deeper still unpublished work by the same author". Compare also the corresponding discussion

the second half of the 1920s. That is to say, within a systematic logical framework of simple type theory, influenced by Whitehead and Russell's *Principia*, he set out to investigate the relationships between the three different notions of completeness suggested by Fraenkel. Carnap's own terms for these notions were "*Entscheidungsdefinitheit*" (deductive completeness), "*Nicht-Gabelbarkeit*" (semantic completeness, Veblen's notion of "non-disjunctive"), and "*Monomorphie*" (categoricity).[83]

The cornerstone of Carnap's work, as reflected in his *Axiomatik*, is a theorem called the "*Gabelbarkeitssatz*". It essentially states that being "nicht-gabelbar" (semantically complete) implies being "monomorph" (categorical).[84] Unfortunately, Carnap's proof of this theorem is faulty, as he eventually came to realize himself. This realization led him to abandon his entire metatheoretic project around 1930. In particular, he decided not to publish the *Axiomatik*, in spite of having already completed a substantial manuscript.[85] Nevertheless, the work was not without immediate influence; for it seems to have served as a catalyst for the thoughts of Carnap's then-student Kurt Gödel, who was one of the few people to have read Carnap's manuscript.

There are several aspects of Carnap's failure in trying to prove the "*Gabelbarkeitssatz*". In particular, he in effect assumed that any consistent theory has a model that is definable within simple type theory, which is false.[86] More generally, he tried to combine a formal axiomatic approach with a genetic logicist standpoint, with the result that he was less than fully clear about the relations among various syntactic and semantic facts and properties. And fundamentally, the work lacks the subsequent sharp distinction between syntax and semantics, between object-language and metalanguage. Despite these flaws, we should recognize as one of Carnap's main contributions in the *Axiomatik* to have explicitly conjectured the "*Gabelbarkeitssatz*", i.e., the claim that semantic completeness of a finite system of axioms implies its categoricity in the context of the simple theory of types.

Another issue that Carnap considered in his investigation—one that was central to the planned, but less finished second part of the *Axiomatik*—

in (Awodey and Carus 2001).

[83] (Carnap 2000), pp. 127ff.

[84] Carnap states the theorem in the contrapositive form: being *polymorph* (non-monomorph) implies being *gabelbar*; *ibid.*, p. 133. Note that the result is mentioned in print already in (Carnap 1927).

[85] Some brief remarks were published in (Carnap 1930a) and (Carnap 1930b).

[86] If this were true, Carnap's proof of the *Gabelbarkeitssatz* would essentially go through. See (Lindenbaum and Tarski 1935), p. 391, Theorem 10.

was, again, the connection between Hilbert's "*Axiom der Vollständigkeit*" and the other three notions of completeness. In this connection, Carnap's central contribution was to note that Hilbert's axiom can be seen as a "extremal axiom", more specifically a "maximality axiom", in that it says that no model can be extended without violating one of the other axioms. As Carnap also noted, the induction axiom of Peano arithmetic can be seen as an analogous "minimality axiom"; it implies that no model can be restricted to a proper subset without violating one of the other axioms. Furthermore, both of these "extremal" axioms lead to categorical, and thus semantically complete, theories. Based on these observations, Carnap raised the further question of how this phenomenon generalizes, and he again arrived at some interesting partial results.[87]

Despite its various shortcomings, Carnap's logical and metatheoretic work from the 1920s—building on that of Fraenkel—remains one of the most systematic treatments of higher-order axiomatics and the relation between categoricity, the various notions of completeness, and line-completeness, specifically in the framework of the simple theory of types. Admittedly, this status is due less to its scope and depth, which is rather limited, than to the subsequent historical shift away from higher-order logic. Influenced by the results of Hilbert, Gödel, and Tarski, much subsequent work has focussed instead on the model theory of first-order logic.[88] Fruitful and important as this has turned out to be, from the perspective of this paper it appears that research into formal axiomatics has been truncated and somewhat disrupted in its progress by the ensuing neglect of higher-order axiomatics.

In Part II below we will now try to suggest how such investigations might proceed, picking up some of the historical threads that have been identified in Part I, and making use of some new methods and results that were not available at the time when these inquiries were dropped.

[87]These were published later in the 1936 paper "*Über Extremalaxiome*", co-written with his student F. Bachmann (see (Carnap and Bachmann 1936), translated as (Carnap and Bachmann 1981)). Compare (Fraenkel and Bar-Hillel 1956), pp. 86–90, for one of the few discussions of the results in that paper.

[88]Cf. (Corcoran 1991): "By the 1930s finite, categorical axiom systems were known for various non-elementary (higher-order) geometrical theories. [...] As certain logicians, including Tarski, came to doubt the foundational significance of higher-order logic, these results seemed to lose some of their importance and to be seen more as challenges to attempt a construction of adequate elementary (first-order) foundations of geometry".

Part II
Recent Developments and Results

4 Higher-order axiomatics

4.1 Limitations of first-order logic

In this part of the paper we are interested in, among other things, the use of logic as a tool in formal axiomatics as discussed in Part I. We occasionally find fault with the standard framework of first-order logic and set-theoretic semantics, and we consider several alternatives. To avoid any misunderstanding, we want to stress here that we are not proposing adopting new foundations for mathematics, and we firmly acknowledge that first-order logic and set-theoretic semantics are important and useful tools in formal axiomatics. Moreover, it is clear that use of second- and higher-order logic involves presuppositions which go beyond those of first-order and that some thinkers find questionable.[89]

That said, the evident difficulty involved in using first-order logic (hereafter FOL) in formal axiomatics is its inability to fully characterize structures with infinite models. The Löwenheim-Skolem theorems show that it is impossible to fully axiomatize an infinite mathematical structure, even up to isomorphism, using only FOL. It follows that FOL is not suitable for characterizing the basic objects of mathematics, like the natural, real, and complex numbers, and the Euclidean spaces.[90]

Moreover, many objects of mathematical study today are described generally by axioms that are not intended to be categorical, but are not of first order either. For example, rings with conditions on ideals, like Noetherian or principal ideal domains; structures on manifolds like vector bundles or tensor fields; the various kinds of spaces used in functional analysis like Hilbert and Banach spaces, and even classical mathematical objects like Euclidean and projective spaces are all determined axiomatically. It is hardly an exaggeration to say that the axiomatic method has succeeded, since its modern beginnings around 1900, in taking over mathematics. But, as the examples just mentioned illustrate, it is not just FOL that is being so widely used.

Of course, one can describe the *models* of such non-first-order axiomatic notions in terms of set theory. But this does not alter the fact that the axiomatic presentation is essentially higher-order. Nor will it do, in such cases,

[89] See (Corcoran 2001) and (Jané 1993) for discussion.

[90] See (Tennant 2000) for another interesting weakness of FOL

to treat higher-order logic as many-sorted first-order logic, as is occasionally suggested. For in specifying structures such as those just mentioned involving higher types of relations or functions, it is essential that these types be interpreted as such, and not as additional first-order structure, if the axiomatization is to serve its intended purpose. We thus believe that higher-order axiomatic theories are best recognized and studied on their own terms, rather than being converted into set theory or first-order logic.

4.2 Higher-order logic

We present here a simple and fairly standard extension of FOL which has the expressive capacity to formulate many of the axiomatic treatments of modern mathematics. Logical languages of this general kind, which are descendant from the type theory mentioned in 3.1 above, are usually called *higher-order logic* or *simple type theory.*[91]

Higher-order systems of logic are those having variables and quantification over "higher types" of relations or functions among the elements of "lower type". Thus, for example, one can extend the usual language $(R, +, \cdot, 0, 1)$ of ring theory by adding also variables $X, Y, \ldots$ ranging over subsets of the domain R. This allows one to axiomatize e.g. principal ideal domains by adding to the theory of commutative rings the familiar condition:

$$\forall I \subseteq R\,(\text{“}I \text{ is an ideal”} \rightarrow \exists x(\, I = (x)\,)) \tag{1}$$

where "I is an ideal" and the principle ideal (x) are defined as usual. Of course, one also adds some logical vocabulary to express subset formation and membership.

We now give an informal description of a particular language of higher-order logic that is sufficient for the purposes of our further discussion. More details of related systems can be found e.g. in (Lambek and Scott 1986).

The language of HOL

The language of higher order logic (HOL) consists of type symbols, terms, and formulas. We write $\tau : X$ to indicate that the term τ has type X.

[91]Type theory is currently experiencing a sort of renaissance because of its applications in computer science. There are literally hundreds of different logical systems that can be called "higher-order logic" or "type theory".

Types In addition to basic type symbols $A, B, \ldots$, and a type P of formulas, further types are built up inductively by the type-forming operations:

$$X \times Y,\ X \to Y,\ P(X)$$

Terms In addition to variables of each type $x_1, x_2, \ldots : X$, and possibly some basic, typed constant symbols, further terms are built up inductively by the term-forming operations:

$$\langle \sigma, \tau \rangle,\ p_1(\tau),\ p_2(\tau)$$
$$\alpha(\tau),\ \lambda x : X.\sigma$$
$$\{x : X \mid \varphi\}$$

Formulas In addition to equations $\sigma = \tau$ and atomic formulas $\tau \in \alpha$, further formulas are built up inductively by the usual logical operations:

$$\neg\varphi,\ \varphi \wedge \psi,\ \varphi \vee \psi,\ \varphi \Rightarrow \psi,\ \forall x : X(\varphi),\ \exists x : X(\varphi)$$

The type of a term is determined in the expected way by the types of the terms used in forming it, and these formations are subject to some obvious conditions for significance; e.g. $p_1(\tau) : A$ if $\tau : A \times B$. We make use of the usual conventions in writing formulas whenever convenient, such as writing $\langle x, y, z \rangle$ for $\langle x, \langle y, z \rangle \rangle$. Note that we have included the possibility of basic type and constant symbols, to be used as the basic language of an axiomatic theory. The theory of rings, for example, has one basic type symbol, say R, and the following basic constant symbols indicated with their types:

$$0,\ 1 : R$$
$$+,\ \cdot : R \times R \to R$$

Generally, we define a *theory* to consist of a basic *language* of type symbols and constants, together with a set of sentences in that language, called the *axioms*. We shall assume here that a theory has finitely many basic symbols and axioms, although there is no reason in principle why one cannot consider infinite theories. In these terms, e.g. the theory of rings thus consists of the language $(R, +, \cdot, 0, 1)$ and the usual handful of axioms for rings (with unit); and the theory of principal ideal domains results by adding the further axiom (1) above.

We emphasize that this use of HOL for presenting axiomatic theories, while familiar enough from everyday mathematical practice, is quite different from the original use intended by logicists like Frege and Russell, and also from that made of it by Carnap in his *Axiomatik* (Carnap 1928), mentioned in 3.2 above. These pioneers had what has been called a "universal" conception of logic ((van Heijenoort 1967),(Goldfarb 1979)), according to which there is a single logical system with a single, fixed domain of quantification (namely, "everything"), and with fixed higher types consisting of "all" functions, concepts, propositional functions, etc. By contrast, the conception in use here has (possibly several) basic types, which can be interpreted in various ways, just as is common in the semantics of first-order logic. Indeed, the clearest way to understand the language of HOL presented here is as an extension of the usual language of FOL by adding higher types and their associated terms, and then building FOL formulas as usual from the terms and variables of those types. In particular, any conventional theory in FOL is also a theory in HOL in the present sense.

When needed, a system of formal deduction can be specified in the usual way, as a formal system with logical axioms and rules of inference. One such system is outlined in the appendix below, but we emphasize that there are many equivalent formulations.

4.3 Semantics

The semantics for HOL is essentially an extension of that for FOL, adjusted to take advantage of the simplifications resulting from the presence of additional types (see e.g. remark 10 below). We shall assume given a "semantic universe" with suitable structure for interpreting the language of HOL. Here we use sets and functions, but later we will generalize to other "universes" (suitable categories) with the required structure.

Rather than stating the formal definition of a model of a theory, we shall give a particular case of it which should be sufficient for the reader to infer the general notion.[92] Suppose we have a theory of the form (A, c, α), with one basic type, one constant, and one axiom. For instance, it might be the theory of semi-groups, with c being $\cdot : A \times A \to A$ and α being the associativity law:

$$\forall x, y, z : A \quad x \cdot (y \cdot z) = (x \cdot y) \cdot z.$$

An *interpretation* assigns to each type X a non-empty[93] set $[\![X]\!]$, in such a way that:

[92] See (Lambek and Scott 1986), (MacLane and Moerdijk 1992), (Awodey and Butz 2000) for details.

[93] This restriction merely simplifies the deductive calculus given in the appendix.

$$\begin{aligned}
[\![A \times B]\!] &= [\![A]\!] \times [\![B]\!] && \text{(the Cartesian product)} \\
[\![A \to B]\!] &= [\![B]\!]^{[\![A]\!]} && \text{(the set of functions)} \\
[\![P(A)]\!] &= \mathcal{P}([\![A]\!]) && \text{(the power set)} \\
[\![P]\!] &= \{\top, \bot\} && \text{(any two-element set)}
\end{aligned}$$

A term $\tau(x) : Y$ containing a free variable $x : X$ is interpreted as a function $[\![\tau]\!] : [\![X]\!] \to [\![Y]\!]$, in such a way that:

$$\begin{aligned}
&[\![c]\!] \in [\![X]\!] && \text{for a basic constant } c : X \\
&[\![x]\!] = 1_{[\![X]\!]} : [\![X]\!] \to [\![X]\!] && \text{(identity function)} \\
& && \text{for a variable } x : X \\
[\![\langle \sigma, \tau \rangle]\!] &= ([\![\sigma]\!], [\![\tau]\!]) && \text{(the ordered pair)} \\
[\![p_i(\tau)]\!] &= \pi_i([\![\tau]\!]) && \text{(the } i\text{-th projection)} \\
[\![\alpha(\tau)]\!] &= [\![\alpha]\!]([\![\tau]\!]) && \text{(functional application)} \\
[\![\lambda x : X.\sigma]\!] &= \text{the function } x \mapsto [\![\sigma]\!] \\
[\![\{x : X \mid \phi\}]\!] &= \text{the subset } \{x \epsilon [\![X]\!] \mid [\![\phi]\!] = \top\} \\
[\![\tau \in \alpha]\!] &= \top \text{ iff } [\![\tau]\!] \epsilon [\![\alpha]\!] \\
[\![\phi \wedge \psi]\!] &= [\![\phi]\!] \wedge [\![\psi]\!] && \text{and similarly for } \neg, \vee, \Rightarrow \\
[\![\forall x : X \phi]\!] &= \top \text{ iff for all } x \epsilon [\![X]\!],\ [\![\phi]\!] = \top \\
[\![\exists x : X \phi]\!] &= \top \text{ iff for some } x \epsilon [\![X]\!],\ [\![\phi]\!] = \top
\end{aligned}$$

Note that we use the boolean operations $\wedge$, etc., on $\{\top, \bot\}$ to interpret the corresponding logical operations.

An interpretation $[\![-]\!]$ *satisfies* a sentence σ ("σ is *true* under $[\![-]\!]$") just if

$$[\![\sigma]\!] = \top$$

Of course, a *model* of a theory is an interpretation that satisfies the axioms. If M is a model, we also write $[\![-]\!]_M$ when considering it as just an interpretation, and we use the notation:

$$M \models \sigma \quad \text{for} \quad [\![\sigma]\!]_M = \top$$

Remark 8. Although it may look a bit unfamiliar at first sight, this definition agrees with the usual one for models of a first-order theory. For instance, in the example above of semi-groups, an interpretation in the present sense consists of a set $[\![A]\!]$ equipped with a binary operation $[\![\cdot]\!] : [\![A]\!] \times [\![A]\!] \to [\![A]\!]$. A model is such a structure for which

$$([\![A]\!], [\![\cdot]\!]) \models \forall x, y, z : A(x \cdot (y \cdot z) = (x \cdot y) \cdot z),$$

which is easily seen to mean just that the operation is associative.

Remark 9. According to our definition, the higher types of functions and relations are interpreted by the corresponding sets (in the conventional terminology, the models are "standard models" rather than "Henkin models"). Observe that such an interpretation is fully determined by the interpretation of the basic language. Thus in particular, an interpretation in this sense of a first-order language is just a first-order structure.

Remark 10. The "internal" notion of satisfaction used here may be unfamiliar; it differs from the more customary, "external" notion of elementary model theory, in that *truth* is represented as an element of a set of truth values $\{\top, \bot\}$, and a formula $\varphi(x)$ where $x : X$ is represented as a function,

$$[\![\varphi(x)]\!] : [\![X]\!] \to \{\top, \bot\}.$$

Of course, $[\![\varphi(x)]\!]$ is just the characteristic function of the subset:

$$[\![\{x : X | \varphi(x)\}]\!] = \{a \in [\![X]\!] \mid [\![X]\!] \models \varphi(a)\} \subseteq [\![X]\!]$$

A sentence (closed formula) is therefore interpreted as one of the truth values $\top$ or $\bot$, with the "true" sentences ($= \top$) being exactly those that hold under the interpretation.

The reason for internalizing truth in this way is that, while it is equivalent to the external approach for set-theoretic semantics, this internal notion can easily be generalized to other semantic universes in a way that external semantics cannot. A similar procedure is sometimes used in connection with *boolean-valued models.*

We now use the semantics to define the notion of *semantic consequence* $\phi \models \psi$ between sentences in the usual way:

$$\phi \models \psi \quad \text{if for every interpretation, } [\![\phi]\!] \leq [\![\psi]\!]$$

where the ordering of truth-values is the usual one, $\bot \leq \top$. The notion of semantic consequence with respect to a theory is defined in the expected way, by considering only those interpretations that are models of the theory.

Remark 11. One sometimes hears it said that HOL is "stronger" than FOL, but this is only so with respect to its expressive capacity, not its semantic consequences. More precisely, the realtion of higher-order semantic consequence is conservative over first-order semantic consequence. For let $\mathbb{T}$ be a first-order theory, regarded as a theory in HOL. The models of $\mathbb{T}$ in the HOL sense are then exactly the models in the usual FOL sense. Thus if the first-order sentence ϕ is true in every HOL model, then it is semantically valid in the sense of FOL.

Remark 12. Semantic consequence for HOL differs from that for FOL in several important respects: it is not compact; the usual Löwenheim-Skolem theorems do not hold; and its theorems are not recursively enumerable (see (Shapiro 1991)).

4.4 Completeness and categoricity

We can now consider more precisely the question of how completeness and categoricity for an axiomatic theory are related in the context of HOL. As already mentioned in 3.2 above, the main early studies are notably (Fraenkel 1928), (Carnap 1928), and (Lindenbaum and Tarski 1983).[94] Briefly, the main positive results are that categoricity implies completeness generally, as for theories in FOL, while in certain cases the converse also holds, which is perhaps more surprising.

Proposition 13. *If a theory $\mathbb{T}$ is categorical, then it is semantically complete.*

Proof. (sketch) Given categorical $\mathbb{T}$, it suffices to show that if $M \models \sigma$ for some model M and sentence σ, then also $N \models \sigma$ for any other model N. But since $\mathbb{T}$ is categorical, there is an isomorphism of $\mathbb{T}$-models $i : M \cong N$, in the usual sense. Now it is easy see that isomorphisms preserve satisfaction, just as in the first-order case. In more detail, one shows by structural induction that for any formula ϕ, one has $[\![\phi]\!]_M = [\![\phi]\!]_N \circ i^n$, as maps $M^n \to \{\top, \bot\}$, where there are n free variables in ϕ, and $i^n : M^n \cong N^n$ is the induced isomorphism on cartesian products. Thus in particular, if $M \models \sigma$ for some sentence σ, then $\top = [\![\sigma]\!]_M = [\![\sigma]\!]_N$, and so also $N \models \sigma$. □

The more interesting question in this connection is, under what conditions does the converse of proposition 13 hold? As already noted above, our

[94]Some recent authors who have also called attention to this topic are (Awodey and Carus 2001), (Read 1997), (Corcoran 1980), and (Corcoran 1981). This section addresses a question raised in the latter.

restriction to finite sets of axioms is essential here. Indeed, it is not hard to find non-isomorphic models that are logically equivalent (since the number of sets of sentences is bounded). The (infinite) "theory" of such a model would then be semantically complete but not categorical.

For the finite theories under consideration here, however, the situation is rather different. As already mentioned in 3.2, in (Carnap 1928) the implication from semantic completeness to categoricity was conjectured and an erroneous proof was offered. The following (correct) proof of a special case is due to Dana Scott:[95]

Proposition 14. *If a theory $\mathbb{T}$ has only one basic type and no basic constant symbols, then $\mathbb{T}$ is categorical if it is semantically complete.*

Proof. (sketch) Let σ be the conjunction of the finitely many axioms, and define the new sentence

$$\begin{aligned}\sigma_0 &=_{df} \sigma \wedge (\forall U : P(X))(\sigma^U \to U \cong X) \\ &= \text{“}X \text{ is the least subset of } X \text{ that satisfies } \sigma\text{”}\end{aligned}$$

in which X is the basic type, U is a variable of type $P(X)$, $U \cong X$ is expressed by the usual definition of isomorphism, and σ^U is a new sentence derived from σ by relativizing all types and quantifiers occurring in σ from X to U.

If σ is satisfied, then so is σ_0 (by the axiom of choice for sets). But if σ is also complete, then we claim that $\sigma \Leftrightarrow \sigma_0$. For if $M \models \sigma$, then we can take some $M' \subseteq M$ such that $M' \models \sigma_0$; since then also $M' \models \sigma$, we also have $M \models \sigma_0$, since σ is complete. But σ_0 is evidently categorical, so σ must also be categorical. □

While it is not difficult to extend this result to a few other cases, we do not know the extent to which it holds in general. Some easy sufficient conditions for the categoricity of a finite theory, given its semantic completeness, are having a definable model (Lindenbaum and Tarski 1983), having a model with no proper submodels, and being categorical in some power. The latter follows from the fact—easily inferred from the foregoing theorem—that all models of a semantically complete theory must have the same cardinality. We know of no counter-examples to the conjecture that semantic completeness of a finite theory implies categoricity in general.

[95] Scott produced this proof in response to a talk on Carnap's failed work by the first author. See also (Awodey and Carus 2001).

In sum, it seems that Carnap's conjecture remains undecided, with little indication as to which way it will go. This is surely to be counted as one of the leading open questions in higher-order axiomatics

5 Topological semantics

In this section we consider an alternative to the usual set-theoretic semantics for HOL.[96] This is drawn from category theory, and is a special case of so-called "topos semantics", which we won't consider in general (see (Lambek and Scott 1986),(MacLane and Moerdijk 1992),and (Awodey and Butz 2000)). The topological semantics outlined here should however suffice to give the reader a general impression of what is involved in interpreting HOL in semantic "universes" other than that of sets.

We first briefly review the motivation for considering alternate semantics for HOL. The first and most obvious reason is that the set-theoretic semantic consequence relation is not deductively axiomatizable in any reasonable sense. Specifically, given a conventional deductive consequence relation $\varphi \vdash \psi$, the Gödel Incompleteness Theorem tells us that this relation cannot be *complete* in the sense of 2.1, Definition 1 with respect to set-theoretic semantic consequence.

This does not necessarily mean that higher-order deduction is somehow defective, however. It is at least sound for set-valued semantics, in the sense that $\varphi \vdash \psi$ implies $\varphi \models \psi$. Moreover, it is conservative over first-order deduction, by a simple argument from the semantic conservativity mentioned in section 4.3 above. And as we shall see below, it is in fact complete with respect to the topological semantics to be considered here.

Another reason to broaden the scope of semantics for HOL is that, like completeness, this also affects the notion of *categoricity* for axiomatic theories, effectively making it a stronger condition. Indeed, since categoricity is a semantic notion, restricting semantics to sets makes it dependent on often nontrivial properties of sets, which can have peculiar, unwanted consequences; in Example 15 below, for instance, we indicate a theory that is categorical just in case the continuum hypothesis holds. The categoricity of certain axiomatic theories like the natural and real numbers seems to provide confirmation of their adequacy, independent of the more subtle properties of sets. Generalizing the range of semantics conforms better to this intention, as will be discussed further in section 6 below.

[96] This section draws on (Awodey 2000), which the reader can consult for more detail.

Finally, one simple reason for considering alternate semantics is that one is interested in the semantic objects. The possibility of using logic to reason about structures on objects other than sets (as happens with e.g. topological groups) makes the systematic investigation of such objects useful in itself. This is indeed the case with the topological semantics considered below; the semantic objects employed (sheaves) are everyday mathematical objects.[97] This is not the case for the most familiar alternate semantics for HOL, the so-called "Henkin models". These are used only for proving deductive completeness, and have no independent mathematical interest.

The objects used in topological semantics are "continuously varying sets", in a sense made precise in 5.2 below. We first motivate this idea in 5.1 by considering an analogy to the ring of continuous, real-valued functions on a topological space. That example also shows how continuous variability can be used to violate some properties of constants, which is essentially what permits the completeness of higher-order deduction with respect to topological semantics, discussed in 5.3.

5.1 Ring of continuous functions

The real numbers $\mathbb{R}$ form a topological space, an abelean group, a commutative ring, a complete ordered field, and much more. Let us consider the properties expressed in just the *language of rings:*

$$0, 1, a + b, a \cdot b, -a$$

and first-order logic. For example, $\mathbb{R}$ is a field:

$$\mathbb{R} \models \forall x(x = 0 \vee \exists y \; x \cdot y = 1)$$

Now consider the *product ring* $\mathbb{R} \times \mathbb{R}$, with elements of the form

$$r = (r_1, r_2)$$

and the *product operations*:

$$\begin{aligned} 0 &= (0, 0) \\ 1 &= (1, 1) \\ (x_1, x_2) + (y_1, y_2) &= (x_1 + y_1, x_2 + y_2) \\ (x_1, x_2) \cdot (y_1, y_2) &= (x_1 \cdot y_1, x_2 \cdot y_2) \\ -(x_1, x_2) &= (-x_1, -x_2) \end{aligned}$$

[97] See e.g. (Hartshorne 1977), (Iversen 1986).

Since these operations are still associative, commutative, and distributive, $\mathbb{R} \times \mathbb{R}$ is still a ring.

But the element $(1,0) \neq 0$ cannot have an inverse, since $(1,0)^{-1}$ would have to be $(1^{-1}, 0^{-1})$. Therefore $\mathbb{R} \times \mathbb{R}$ *is not a field.*

In a similar way, one can form the more general product rings $\mathbb{R} \times \ldots \times \mathbb{R} = \mathbb{R}^n$, or $\mathbb{R}^I$ for any index-set I.

While not in general fields, product rings $\mathbb{R}^I$ are always (von Neumann) *regular*:

$$\mathbb{R}^I \models \forall x \exists y \ (x \cdot y \cdot x = x).$$

For, given x, we can take $y = (y_i)$ with:

$$y_i = \begin{cases} x_i^{-1}, & \text{if } x_i \neq 0 \\ 0, & \text{if } x_i = 0 \end{cases}$$

One can produce rings that violate even more properties of $\mathbb{R}$ by passing to "continuously varying reals". But what is a "continuously varying real number"?

Let X be a topological space. A "real number r_x varying continuously over X" is just a continuous function:

$$r : X \to \mathbb{R}$$

We equip these functions with the pointwise operations:

$$(f+g)(x) = f(x) + g(x), \quad \text{etc.}$$

The set $\mathcal{C}(X)$ of all such functions then forms a *subring* of the product ring over the index set $|X|$ of points of the space X, $\mathcal{C}(X) \subseteq \mathbb{R}^{|X|}$. But unlike the product ring, $\mathcal{C}(X)$ is not regular:

$$\mathcal{C}(X) \not\models \forall f \exists g \ (f \cdot g \cdot f = f)$$

For take e.g. $X = \mathbb{R}$ and $f(x) = x^2$, then we must have:

$$g(x) = \frac{1}{x^2}, \qquad \text{if } x \neq 0$$

but of course:

$$g(0) = \lim_{x \to 0} g(x) = \lim_{x \to 0} \frac{1}{x^2} = \infty$$

so there can be no *continuous* g satisfying $f \cdot g \cdot f = f$.

Thus the "continuously varying reals" $\mathcal{C}(X)$ have even fewer properties of the field of "constant" reals $\mathbb{R}$ than do the product rings $\mathbb{R}^I$. In this way, passing from constants to continuous variation "abstracts away" some properties of the constants.

5.2 Continuously variable sets

Just as the real numbers could be generalized to the "continuously variable reals" (continuous functions), we now generalize the notion of a set to that of a "continuously variable set," i.e. a *sheaf.*

As a first step, observe that the type-forming operations of product, powerset, equality, etc. can be interpreted in other "universes" of sets. Indeed, consider in the universe of "pairs of sets," **Sets** $\times$ **Sets**. The objects have the form:

$$A = (A_1, A_2)$$

and the operations are defined componentwise:

$$\begin{aligned} (A_1, A_2) \times (B_1, B_2) &= (A_1 \times B_1, A_2 \times B_2) \\ \mathcal{P}(A_1, A_2) &= (\mathcal{P}(A_1), \mathcal{P}(A_2)) \\ P &= (P, P) \end{aligned}$$

Term-formation is similarly componentwise. Indeed, the logical operations can also be defined componentwise:

$$\begin{aligned} (a_1, a_2) \in (A_1, A_2) &= (a_1 \in A_1, a_2 \in A_2) \\ (\varphi_1, \varphi_2) \wedge (\psi_1, \psi_2) &= (\varphi_1 \wedge \psi_1, \varphi_2 \wedge \psi_2) \\ &\text{etc.} \end{aligned}$$

This interpretation of the logical language models HOL in the sense that the usual logical axioms and rules of inference (e.g. as given in the appendix) are all validated. On the other hand, it does not satisfy *all* the properties of **Sets**. For example:

$$\mathbf{Sets} \models A \cong 0 \vee \exists x \; x \in A$$

But in **Sets**$\times$**Sets** we can take as A the object $(1, 0)$, which is not isomorphic to 0, and then $a \in (1, 0)$ means $a = (a_1, a_2)$ with $a_1 \in 1$ and $a_2 \in 0$, which is impossible.

Just as in the case of rings, we can also generalize to **Sets** $\times \ldots \times$ **Sets** $=$ **Sets**^{n}, and indeed to **Sets**^{I} for any index set I, to get the "universe" of I-indexed families of sets.

All such "product universes" have some things in common, e.g. they all satisfy the axiom of choice (which, by the way, can be seen to be formally analogous to regularity for rings). To find even more general "universes" we consider even more general families of sets,

$$(F_x)_{x \in X}$$

varying continuously over an arbitrary topological space X. But what should a "continuously varying set" be? The problem is that we cannot simply take a "continuous set-valued function"

$$F : X \to \mathbf{Sets}$$

as we did for rings of real-valued functions, since **Sets** is not a topological space.

In modern mathematics, one often encounters continuously varying structures; let us recall how this is typically done, in order to find the notion we seek. A "continuously varying space" $(Y_x)_{x \in X}$ over a space X is called a *fiber bundle*. It consists of a space $Y = \sum_{x \in X} Y_x$ and a continuous "indexing" projection $\pi : Y \to X$, with $\pi^{-1}\{x\} = Y_x$, as indicated below.

$$\begin{array}{c} Y = \sum_{x \in X} Y_x \\ \big\downarrow \pi \\ X \end{array}$$

A "continuously varying group" $(A_x)_{x \in X}$ is a *sheaf of groups*. It consists essentially of a fiber bundle $\pi : A = \sum_{x \in X} A_x \to X$ satisfying the additional requirements:

1. each A_x is a group,
2. the operations in the fibers A_x "fit together continuously",
3. π is a local homeomorphism (see below).

We can answer the question of what a "continuously varying set" should be: it is a *sheaf of sets*, i.e. a fiber bundle,

$$\begin{array}{c} F = \sum_{x \in X} F_x \\ \big\downarrow \pi \\ X \end{array}$$

such that π is a *local homeomorphism*, in the sense that each point $y \in F$ has some neighborhood U on which π is a homeomorphism $U \xrightarrow{\sim} \pi(U)$.

This ensures in particular that each fiber $F_x = \pi^{-1}(x)$ is discrete, and that the variation across the fibers is continuous, in a suitable sense.

To define the semantics of HOL in sheaves, one needs to specify the basic type-forming and logical operations. Some of these can be defined pointwise, $(F \times G)_x \cong (F_x \times G_x)$. Others, however, cannot; for instance, the exponential G^F of sheaves F, G is the "sheaf-valued hom" $\hom(F, G)$, defined in terms of *germs of continuous maps* $F \to G$, for which $(G^F)_x \ncong G_x^{F_x}$. This is what makes topological semantics different from the product semantics of indexed families.

Like the product universes, the universe $\mathrm{sh}(X)$ of all sheaves on a given space X models HOL in the sense that the axioms are all true and the rules of inference are all sound. But in general sheaves violate the axiom of choice. Indeed, one can find sheaf models of HOL that also violate many other properties of sets.

5.3 Topological completeness

If we think of sheaves as sets varying continuously in a parameter, the constant sets occur as the special case of no variation. The semantics given in 4.3 above apply *mutatis mutandis* to yield topological semantics, with standard set-theoretic semantics as a special case.

Some logical statements that are not true of variable sets in general are true of all constant sets, as a result of their special properties. In this sense, the logic of the constant sets is quite strong, while the logic of variable sets is much weaker. That is, fewer things are true of all variable sets than are true of constant ones. This is just like the difference between the field of real numbers and the ring of real-valued functions. Now one can ask, what is the logic of continuously varying sets? That is to say, which sentences of HOL are true in all sheaf models? The answer is given by the following theorem from (Awodey and Butz 2000).[98]

Theorem. *HOL is complete with respect to topological semantics.*

The completeness referred to is deductive completeness in the sense of our Definition I.2.1, with respect to the standard, classical deductive consequence relation, as specified in the appendix. Thus if a sentence is true in all topological models, then it is provable.

The reader may wonder how this result is to be reconciled with the Gödel incompleteness of deductive higher-order logic. Roughly speaking,

[98] The proof uses recent results in topos theory (Butz and Moerdijk 1999) that are rooted in geometry.

the situation is this: the sense in which a sentence is "true but unprovable" in Gödel's theorem involves only "true of all *constant* sets," but not "true of all *variable* sets." Thus a "true but unprovable" Gödel-style sentence is only true of the constant sets, but it is violated by some variable ones (else it would be provable).

6 Notions of categoricity

Having now considered completeness and categoricity with respect to standard, set-theoretic semantics, and deductive completeness with respect to alternate semantics, we turn to possible alternate notions of categoricity. We have already mentioned that some axiomatic theories in higher-order logic are categorical in the usual sense that any two (standard) models are isomorphic just in case the sets used to model them are assumed to have certain properties, such as satisfying the continuum hypothesis or the axiom of choice. This is essentially because these properties of sets are expressible in HOL. For example, the following simple theory is categorical just if the continuum hypothesis holds.

Example 15. The theory $\mathbb{T}_0$ has one basic type symbol U, one relation symbol $R : PP(U)$, and two axioms expressing the conditions "U is countably infinite" and "$|U| \precneqq |R|$".

But the idea of categoricity as a basic criterion of adequacy for a system of axioms seems to presume that it is not sensitive to such questions as whether the continuum hypothesis holds. Indeed such issues seem irrelevant to the categoricity of descriptions of at least some classical mathematical notions, like the natural numbers. As was seen in part I above, some version of categoricity was one of the main early conditions of adequacy for axiom systems, quite independently of a precise specification of the theory of sets, or any understanding of their more subtle properties.

In this section, we consider several strengthenings of the notion of categoricity that are not sensitive in this way to special properties of the semantics, although they do have their own peculiarities. The notions considered are called *unique*, *variable*, and *provable* categoricity. Some of the classical theories of greatest interest do indeed have these stronger properties. Finally, we consider the category-theoretic concept of *universality* and its relation to axiomatic descriptions.

6.1 Unique categoricity

This notion strengthens conventional categoricity by requiring that any two models M and N be isomorphic via a *unique* isomorphism $M \cong N$. This is clearly equivalent to saying that the theory at issue is categorical and, furthermore, that its models have no non-trivial automorphisms.[99]

The classical axiomatizations for the natural and real numbers do indeed have this stronger property (the complex numbers do, too, if one eliminates complex conjugation as an automorphism by adding a constant symbol for i). As will become more clear below, axiomatizations are sometimes categorical because there are some natural or canonical maps between models (as opposed to ones gotten, say, by the axiom of choice), and the axioms then suffice to make these canonical maps isomorphisms. The property of unique categoricity also seems to accompany some of the other strengthenings to be considered, and it is found in connection with the category theoretical notion of universality.

6.2 Variable categoricity

We have already considered the notion of a continuously varying model M over a space of parameters X, as made precise by the concept of a *sheaf of models*, which is a model in the "universe" $sh(X)$ of continuously variable sets (cf. 5.2 above). The notion of *variable categoricity* is simply the obvious generalization of categoricity to such variable models:

Definition 16. A theory $\mathbb{T}$ is called *variably categorical* if any two continuously variable models M, N over any space X are isomorphic.

Remarks. 1. This condition requires more than just that there is an isomorphism

$$h_x : M_x \xrightarrow{\sim} N_x \qquad \text{for each } x \in X.$$

In addition, the h_x must fit together to form a single, continuous isomorphism

$$h : M \xrightarrow{\sim} N \qquad \text{over } X.$$

Thus, in effect, the h_x must also vary continuously with the parameter x.

[99] Cf. (Tarski 1983), p. 313.

2. Note that this notion does generalize conventional categoricity, since the conventional notion is the special case of variation over a one-point parameter space. In this sense, conventional categoricity is the limiting or trivial case of variable categoricity.

3. There is, of course, no sense to requiring that models over *different* spaces be isomorphic, since there is no notion of a map between such models (at least in the current situation).

4. There is an obvious *unique* version of this notion, obtained by requiring unique isomorphisms between models.

5. The classical theories of $\mathbb{N}$ and $\mathbb{R}$ have this property — indeed they are *uniquely* variably categorical. The contrived theory $\mathbb{T}_0$ above does not have it, however (even assuming CH). The reason why is roughly that, in a given model M, a variable subset $R_M \subseteq P(U_M)$ might be *pointwise* isomorphic to $P(U_M)$, just for cardinality reasons, without there being a *continuous* isomorphism $R_M \xrightarrow{\sim} P(U_M)$ over the space of parameters X.

As suggested by the last remark above, the basic idea behind variable categoricity is that the strong requirement that the isomorphisms must also be continuously parametrized with the models tends to "break up" accidental or arbitrary choices of maps, and restrict to those that are somehow intrinsic to or canonically associated with the structure at issue. The following notion provides another, rather different, way of restricting the possible isomorphisms, namely by requiring them to be definable or provable.

6.3 Provable categoricity

We want to formulate the idea that the connecting isomorphism between any two models of a categorical theory is definable from the language of the theory, and is provably an isomorphism from the axioms of the theory.[100] To specify this notion, suppose our theory $\mathbb{T}$ is of the form:

$$U,\ f : T(U),\ \alpha(U, f)$$

where U is a basic type symbol, f a basic constant symbol of type $T(U)$, and $\alpha(U, f)$ a sentence in the language U, f and higher-order logic. Here we display U in the type symbol $T(U)$ to remind ourselves that the type of f may contain U as a parameter, e.g. if f represents a binary operation on U,

[100] Cf. (Tarski 1983), p. 310.

then $T(U)$ is $U \times U \to U$. Similarly, the axiom $\alpha(U, f)$ likely contains the basic language U, f.

Now consider the new theory $\mathbb{T}^2$, which is essentially two copies of $\mathbb{T}$ written side-by-side. It has:

- basic types: $U_1,\ U_2$
- basic terms: $f_1 : T(U_1),\ f_2 : T(U_2)$
- axioms: $\alpha_1(U_1, f_1),\ \alpha_2(U_2, f_2)$

where $T(U_1)$ is built from U_1 in the same way that $T(U)$ was built from U, e.g. if $f : U \times U \to U$, then $f_1 : U_1 \times U_1 \to U_1$, and similarly for $T(U_2)$. The axioms are similarly just the axiom α of $\mathbb{T}$ with the respective substitutions of (U_1, f_1) and (U_2, f_2) for (U, f).

Observe that a model of $\mathbb{T}^2$ is just a pair of models of $\mathbb{T}$,

$$\mathbf{Mod}(\mathbb{T}^2) = \mathbf{Mod}(\mathbb{T}) \times \mathbf{Mod}(\mathbb{T}).$$

Definition 17. $\mathbb{T}$ is called *provably categorical* if:

$$\mathbb{T}^2 \vdash \exists h : U_1 \to U_2 \text{ ``h is a $\mathbb{T}$-model isomorphism''}$$

where the formula "h is a $\mathbb{T}$-model isomorphism" is to be spelled out in higher-order logic in the obvious way.

The idea behind provable categoricity is that the theory $\mathbb{T}$ has enough "logical strength" on its own to ensure that any two $\mathbb{T}$-models are isomorphic.

Remarks. 1. This notion is plainly dependent on the logical consequence relation represented by $\vdash$. Here we are assuming the classical, syntactic consequence relation in higher-order logic (as given in the Appendix). A different (weaker) notion results if we take instead e.g. semantic consequence for classical **Set**-valued semantics. That notion is clearly equivalent to conventional categoricity. Of course, any theory that is provably categorical is also categorical.

2. A stronger condition results from a weaker notion of logical consequence $\vdash$. For example, using intuitionistic provability instead of classical by omitting the law of excluded middle makes it more difficult for a theory to be provably categorical. It is not hard to make up theories that are provably categorical classically, but not so intuitionistically.

3. The familiar theories of natural and real numbers are provably categorical (even intuitionistically). The contrived theory $\mathbb{T}_0$ which depends on the continuum hypothesis is evidently not (else one could prove CH in higher-order logic).

As these remarks make clear, there is a connection between provable categoricity and semantic considerations like completeness. Indeed the completeness of the higher-order deductive consequence relation with respect to topological semantics is used in the proof of the following:

Theorem 18. *A theory is provably categorical if and only if it is variably categorical.*

The even stronger notion of *intuitionistically provable categoricity* is equivalent to a certain semantic notion that is phrased in terms of arbitrary *toposes*, but we have chosen not to go into that here.

6.4 Universality

Category theory provides a notion of "unique specification" that is related to categoricity in an interesting way, which remains to be clarified. Although this is not the place for a thorough discussion, it seems at least worth mentioning the basic connection and a couple of examples.

The basic concept we have in mind is that of a *universal mapping property*, which can be used to characterize a particular mathematical structure. The connection with the present topic results from the fact that universal mapping properties are unique characterizations up to isomorphism; any two structures that satisfy a universal mapping property are necessarily isomorphic. Indeed, such structures are *uniquely* isomorphic; so universal mapping properties may be compared with uniquely categorical theories.

The two notions do not seem to be equivalent, however. While some concepts can be formulated both in terms of a categorical, axiomatic theory and a universal mapping property, some concepts seem to be given most naturally in one way or the other, as the following examples illustrate.

Examples. 1. The natural numbers are characterized by the universal mapping property called "natural numbers object", due to Lawvere (Lawvere 1969). In any category with a terminal object 1, consider arbitrary structures of the form:

$$1 \xrightarrow{\quad a \quad} U \xrightarrow{\quad f \quad} U$$

(no conditions on f). A *natural numbers object* is a *universal* structure of this type. That is, one (N, o, s) such that given any such (U, a, f) there is a unique homomorphism $h : (N, o, s) \to (U, a, f)$, i.e. a map $h : N \to U$ such that $ho = a$ and $hs = fh$, as indicated in the commutative diagram below.

$$\begin{array}{ccccc} 1 & \xrightarrow{\;o\;} & N & \xrightarrow{\;s\;} & N \\ \| & & \downarrow{\scriptstyle h} & & \downarrow{\scriptstyle h} \\ 1 & \xrightarrow{\;a\;} & U & \xrightarrow{\;f\;} & U \end{array}$$

This characterization is equivalent to the familiar Peano axioms in categories like **Sets**. It is worth mentioning that it also applies in much more general categories than **Sets**, where the Peano axioms cannot be interpreted.

2. A notion that can be given by a universal mapping property, but not by any familiar axioms, is that of the *free group* on a set of generators. Consider the case of two generators: the *free group* $F(x, y)$ on the elements x, y has the property that for any group G and elements $g, g' \in G$, there is a unique homomorphism $h : F(x, y) \to G$ with $h(x) = g$ and $h(y) = g'$. The concept of a *polynomial ring* is defined by a similar universal mapping property.

3. The real numbers provide an example of a (uniquely) categorical concept that is not determined by any known universal mapping property.

Of course, it may be that one can find axioms for free groups, polynomial rings, etc., or even for any particular universal mapping property, or that the real numbers can be characterized by a suitable universal mapping property.[101] We don't know whether this is the case, but simply mention the connection between categoricity and universality as a direction for possible further research. Indeed, this line of thought seems to be quite closely related to Carnap's work on extremal axioms and Hilbert's Axiom of Line Completeness, mentioned in section I.3.2.

[101] See (Pavlovic and Pratt 1999), (Escardo and Simpson 2001) for some recent attempts.

Appendix: Deduction for higher-order logic

The *deductive consequence relation* $\varphi \vdash \psi$ between formulas is specified by a deductive calculus in the usual way. The following rules of inference could be reduced considerably by defining some logical operations in terms of others. See (Lambek and Scott 1986) for some alternatives.

1. *Order*
   (a) $\varphi \vdash \varphi$
   (b) $\varphi \vdash \psi$ and $\psi \vdash \vartheta$ implies $\varphi \vdash \vartheta$
   (c) $\varphi \vdash \psi$ implies $\varphi[\tau/x] \vdash \psi[\tau/x]$
2. *Equality*
   (a) $\top \vdash \tau = \tau$
   (b) $\tau = \tau' \vdash \varphi[\tau/x] \Rightarrow \varphi[\tau'/x]$
   (c) $\vartheta \vdash \varphi \Rightarrow \psi$ and $\vartheta \vdash \psi \Rightarrow \varphi$ implies $\vartheta \vdash \varphi = \psi$
   (d) $\forall x.(\alpha(x) = \beta(x)) \vdash \alpha = \beta$
3. *Products*
   (a) $\top \vdash \langle p_1\tau, p_2\tau \rangle = \tau$
   (b) $\top \vdash p_i\langle \tau_1, \tau_2 \rangle = \tau_i, \quad i = 1, 2$
4. *Exponents*
   (a) $\top \vdash (\lambda x.\tau)(x) = \tau$
   (b) $\top \vdash \lambda x.\alpha(x) = \alpha \quad (x \text{ not free in } \alpha)$
5. *Elementary logic*
   (a) $\bot \vdash \varphi$
   (b) $\varphi \vdash \top$
   (c) $\neg\neg\varphi \vdash \varphi$
   (d) $\vartheta \vdash \neg\varphi$ iff $\vartheta \wedge \varphi \vdash \bot$
   (e) $\vartheta \vdash \varphi$ and $\vartheta \vdash \psi$ iff $\vartheta \vdash \varphi \wedge \psi$
   (f) $\vartheta \vee \varphi \vdash \psi$ iff $\vartheta \vdash \psi$ and $\varphi \vdash \psi$
   (g) $\vartheta \wedge \varphi \vdash \psi$ iff $\vartheta \vdash \varphi \Rightarrow \psi$

(h) $\vartheta \vdash \varphi(x)$ iff $\vartheta \vdash \forall x \varphi(x)$ (x not free in ϑ)

(i) $\exists x \varphi(x) \vdash \vartheta$ iff $\varphi(x) \vdash \vartheta$ (x not free in ϑ)

The τ's are any terms; φ, ψ, ϑ are any formulas; α, β are any terms of the same exponential type. Substitution $\varphi[\tau/y]$ must include a convention to avoid binding free variables in τ. The type $P(X)$ and the associated terms $\tau \in \alpha$ and $\{x : X|\varphi\}$ are treated as alternate notation for $X \to P$, $\alpha(\tau)$, and $\lambda x : X.\varphi$, respectively.

Acknowledgments. We would like to thank Jeremy Avigad, Andre Carus, Dirk Schlimm, and, especially, John Corcoran, Thomas Scanlan, and George Weaver for very helpful comments on earlier versions of this paper. Parts of it were presented at the PSA Meeting, Vancouver, November 2000 (by Awodey), the Center for Philosophy of Science, University of Pittsburgh, January 2001 (by Awodey), and the APA/ASL Meeting, Central Division, Minneapolis, May 2001 (by Reck). We would also like to thank various members of the audience at these occasions for their questions and their interest in the topic.

References

Aspray, W. (1991). Oswald Veblen and the origins of mathematical logic at Princeton. In T. Drucker (Ed.), *Perspectives on the History of Mathematical Logic*, pp. 54–70. Basel: Birkhäuser.

Awodey, S. (2000). Continuity and logical completeness. Technical Report CMU-PHIL-116, Carnegie Mellon University.

Awodey, S. and C. Butz (2000). Topological completeness for higher-order logic. *Journal of Symbolic Logic 65*, 1168–82.

Awodey, S. and A. Carus (2001). Carnap, completeness and categoricity: The *Gabelbarkeitssatz* of 1928. *Erkenntnis 54*, 145–72.

Bernays, P. (1918). Beiträge zur axiomatischen Behandlung des Logik-Kalküls. Habilitationsschrift, University of Göttingen, unpublished.

Blanchette, P. (1996). Frege and Hilbert on consistency. *Journal of Philosophy 93*, 317–36.

Butz, C. and I. Moerdijk (1999). Topological representation of sheaf cohomology of sites. *Compositio Mathematica 118*, 217–33.

Carnap, R. (1927). Eigentliche und uneigentliche Begriffe. *Symposion 1*, 355–74.

Carnap, R. (1928). Untersuchungen zur Allgemeinen Axiomatik. Manuscript, Carnap Nachlaß, University of Pittsburgh.

Carnap, R. (1929). *Abriß der Logistik.* Vienna: Springer.

Carnap, R. (1930a). Bericht über Untersuchungen zur allgemeinen Axiomatik. *Erkenntnis 1*, 303–10.

Carnap, R. (1930b). Die alte und die neue Logik. *Erkenntnis 1*, 12–26.

Carnap, R. (2000). *Untersuchungen zur Allgemeinen Axiomatik.* Darmstadt: Wissenschaftliche Buchgesellschaft. Edited by T. Bonk and J. Mosterin.

Carnap, R. and F. Bachmann (1936). Über Extremalaxiome. *Erkenntnis 6*, 166–88.

Carnap, R. and F. Bachmann (1981). On extremal axioms. *History and Philosophy of Logic 2*, 67–85. H. G. Bohnert, trans.

Church, A. (1940). A formulation of the simple theory of types. *Journal of Symbolic Logic 5*, 56–68.

Church, A. (1956). *Introduction to Mathematical Logic*, Volume 1. Princeton: Princeton University Press.

Chwistek, L. (1921). Antynomje logiki formalnej. *Przeglad Filozoficzny 24*, 164–71.

Chwistek, L. (1924–1925). The theory of constructive types. *Annales de la Societe Polonaise de Mathematique 2-3*, 9–48 and 92–141.

Chwistek, L. (1967). Antinomies of formal logic. In S. McCall (Ed.), *Polish Logic*, pp. 338–45. Oxford: Clarendon Press. Z. Jordan, trans.

Corcoran, J. (1972). Review of *Reid, C.: Hilbert. Philosophy of Science 39*, 106–8.

Corcoran, J. (1980). Categoricity. *History and Philosophy of Logic 1*, 187–207.

Corcoran, J. (1981). From categoricity to completeness. *History and Philosophy of Logic 2*, 113–19.

Corcoran, J. (1991). Review of *Tarski, A. Collected Papers. Vol. 4. Mathematical Reviews 91.* Review 91h:01104.

Corcoran, J. (1995). Axiomatic method. In R. Audi (Ed.), *The Cambridge Dictionary of Philosophy*, pp. 57–58. Cambridge: Cambridge University Press.

Corcoran, J. (2001). Second-order logic. In C. A. Anderson (Ed.), *Church Memorial Volume*. Dondrecht: Kluwer. To appear.

Corcoran, J., M. Frank, and M. Maloney (1974). String theory. *Journal of Symbolic Logic 39*, 625–37.

DaSilva, J. (2000). Husserl's two notions of completeness. *Synthese 125*, 417–38.

Dedekind, R. (1872). Stetigkeit und irrationale Zahlen. Reprinted in *Gesammelte Mathematische Werke*, Volume 3, R. Fricke et al. (Eds.), Braunschweig: Vieweg, 1932, pp. 315–34.

Dedekind, R. (1888). Was sind und was sollen die Zahlen? Reprinted in *Gesammelte Mathematische Werke*, Volume 3, R. Fricke et al. (Eds.), Braunschweig: Vieweg, 1932, pp. 335–91.

Dedekind, R. (1890). Letter to Keferstein. Published in *From Frege to Gödel*, J. van Heijenoort (Ed.), Cambridge, Mass.: Harvard University Press, 1969, pp. 98–103.

Dedekind, R. (1963). *Essays on the Theory of Numbers*. New York: Dover. W. W. Behman, trans.

Dreben, B. and J. Heijenoort (1986). Introductory note to 1929, 1930 and 1930a. In S. Feferman et al. (Eds.), *Kurt Gödel: Collected Works, Volume 1*, pp. 44–59. New York: Oxford University Press.

Ehrlich, P. (1995). Hahn's *Über die Nichtarchimedischen Größensysteme* and the development of the modern theory of magnitudes and numbers to measure them. In J. Hintikka (Ed.), *From Dedekind to Gödel*, pp. 165–213. Dordrecht: Kluwer.

Ehrlich, P. (1997). From completeness to Archimedean completeness. *Synthese 110*, 57–76.

Escardo, M. and A. Simpson (2001). A universal characterization of the closed interval. In *Proceedings of the 16th Annual Symposium on Logic in Computer Science (LICS'01)*. To appear.

Fraenkel, A. (1923). *Einleitung in die Mengenlehre* (second, revised ed.). Berlin: Springer. First edition published in 1919.

Fraenkel, A. (1928). *Einleitung in die Mengenlehre* (third, revised ed.). Berlin: Springer.

Fraenkel, A. and Y. Bar-Hillel (1956). *Foundations of Set Theory.* Amsterdam: North Holland.

Frege, G. (1879). *Begriffsschrift.* Reprinted by Wissenschaftliche Buchgesellschaft, Darmstadt, 1977.

Gödel, K. (1929). *Über die Vollständigkeit des Logikkalküls.* Ph. D. thesis, University of Vienna. Published in *Collected Works*, Volume 1, S. Feferman et al. (Eds.), Oxford: Oxford University Press, 1986, pp. 60-100.

Gödel, K. (1930). Die Vollständigkeit der Axiome des logischen Funktionenkalküls. Reprinted in *Collected Works*, Volume 1, S. Feferman et al. (Eds.), Oxford: Oxford University Press, 1986, pp. 102-23.

Gödel, K. (1931). Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I. Reprinted in *Collected Works*, Volume 1, S. Feferman et al. (Eds.), Oxford: Oxford University Press, 1986, pp. 144–94.

Goldfarb, W. (1971). Editor's notes. In W. Goldfarb (Ed.), *Jacques Herbrand: Logical Writings.* Dordrecht: Reidel.

Goldfarb, W. (1979). Logic in the twenties: the nature of the quantifier. *Journal of Symbolic Logic 44*, 351–68.

Goldfarb, W. (2001). Frege's conception of logic. In J. Floyd and S. Shieh (Eds.), *Future Pasts: The Analytic Tradition in Twentieth Century Philosophy.* New York: Oxford University Press. To appear.

Hartshorne, R. (1977). *Algebraic Geometry.* New York: Springer.

Henkin, L. (1950). Completeness in the theory of types. *Journal of Symbolic Logic 15*, 81–91.

Hilbert, D. (1899). *Grundlagen der Geometrie* (first German ed.). Leipzig: Teubner.

Hilbert, D. (1900). Über den Zahlbegriff. *Jahresbericht der Deutschen Mathematiker-Vereinigung 8*, 180–84.

Hilbert, D. (1902). *Foundations of Geometry* (first English ed.). LaSalle: Open Court. L. Unger, trans.

Hilbert, D. (1903). *Grundlagen der Geometrie* (second German ed.). Leipzig: Teubner.

Hilbert, D. (1929). Probleme der Grundlegung der Mathematik. *Mathematische Annalen 102*, 1–9.

Hilbert, D. (1971). *Foundations of Geometry* (second English ed.). LaSalle: Open Court. Based on the tenth German edition.

Hilbert, D. (1996). On the concept of number. In W. Ewald (Ed.), *From Kant to Hilbert*, Volume 2, pp. 1089–95. New York: Oxford University Press. W. Ewald, trans.

Hilbert, D. and W. Ackermann (1928). *Grundzüge der theoretischen Logik*. Berlin: Springer.

Hilbert, D. and P. Bernays (1934). *Grundlagen der Mathematik*, Volume 1. Berlin: Springer.

Hodges, W. (1986). Truth in a structure. *Proceedings of the Aristotelian Society, New Series 86*, 135–51.

Huntington, E. V. (1902). A complete set of postulates for the theory of absolute continuous magnitude. *Transactions of the American Mathematical Society 3*, 264–79.

Huntington, E. V. (1903). Complete set of postulates for the theory of real quantities. *Transactions of the American Mathematical Society 4*, 358–70.

Huntington, E. V. (1911). The fundamental propositions of algebra. In J. W. Young (Ed.), *Monographs on Topics of Modern Mathematics*, pp. 151–207. Reprinted by Dover, New York, 1955.

Huntington, E. V. (1917). *The Continuum and Other Types of Serial Order* (second ed.). Cambridge, Mass.: Harvard University Press. Reprinted by Dover, New York, 1955.

Iversen, B. (1986). *Cohomology of Sheaves*. New York: Springer.

Jané, I. (1993). A critical appraisal of second-order logic. *History and Philosophy of Logic 14*, 67–86.

Lambek, J. and P. J. Scott (1986). *Introduction to Higher-Order Categorical Logic*. Cambridge: Cambridge University Press.

Lawvere, F. (1969). Adjointness in foundations. *Dialectica 23*, 281–96.

Lindenbaum, A. and A. Tarski (1934–1935). Über die Beschränkheit der Ausdrucksmittel deduktiver Theorien. In *Ergebnisse eines Mathematischen Kolloquiums, fascicule 7*, pp. 15–22.

Lindenbaum, A. and A. Tarski (1983). On the limitations of the means of expression of deductive theories. In J. Corcoran (Ed.), *Alfred Tarski: Logic, Semantics, Metamathematics*, pp. 384–92. Indianapolis: Hackett. J. H. Woodger, trans.

MacLane, S. and I. Moerdijk (1992). *Sheaves in Geometry and Logic: A First Introduction to Topos Theory.* New York: Springer.

Majer, U. (1997). Husserl and Hilbert on completeness. *Synthese 110*, 37–56.

Mancosu, P. (1998). Introduction and notes. In P. Mancosu (Ed.), *From Brouwer to Hilbert.* New York: Oxford University Press.

Moore, G. (1980). Beyond first-order logic: The historical interplay between mathematical logic and axiomatic set theory. *History and Philosophy of Logic 1*, 95–137.

Moore, G. (1988). The emergence of first-order logic. In W. Aspray and P. Kitcher (Eds.), *History and Philosophy of Modern Mathematics*, pp. 95–135. Minneapolis: University of Minnesota Press.

Pavlovic, D. and V. Pratt (1999). On coalgebra of real numbers. *Electronic Notes in Theoretical Computer Science 19.*

Peano, G. (1889). *Arithmetices Principia, Nova Methodo Exposita.* Turin: Bocca.

Peano, G. (1973). The principles of arithmetic, presented by a new method. In H. Kennedy (Ed.), *Selected Works of Guiseppe Peano*, pp. 101–34. Toronto: University of Toronto Press. H. Kennedy, trans.

Post, E. (1921). Introduction to a general theory of elementary propositions. *American Journal of Mathematics 43*, 163–85.

Ramsey, F. (1926). Mathematical logic. *The Mathematical Gazette 13*, 185–94.

Read, S. (1997). Completeness and categoricity: Frege, Gödel and model theory. *History and Philosophy of Logic 18*, 79–93.

Scanlan, M. (1991). Who were the American postulate theorists? *Journal of Symbolic Logic 56*, 981–1002.

Shapiro, S. (1991). *Foundations Without Foundationalism: A Case for Second-Order Logic.* Oxford: Oxford University Press.

Sieg, W. (1999). Hilbert's program: 1917–1922. *Bulletin of Symbolic Logic 5*, 1–44.

Tait, W. W. (1997). Cantor versus Frege and Dedekind: On the concept of number. In W. W. Tait (Ed.), *Early Analytic Philosophy*, pp. 213–48. Chicago: Open Court.

Tarski, A. (1951). *A Decision Method for Elementary Algebra and Geometry* (second, revised ed.). Berkeley: University of California Press. Prepared for publication with the assistance of J. McKinsey. First edition published by RAND Corporation, Santa Monica, 1948.

Tarski, A. (1983). *Logic, Semantics, Metamathematics* (second ed.). Indianapolis: Hackett. Edited, with introduction, by J. Corcoran; J. H. Woodger, trans. First edition published by Oxford University Press, Oxford, 1956.

Tarski, A. and A. Lindenbaum (1926). Sur l'indépendance des notions primitives dans les systèmes mathématiques. *Annales de la Societe Polonaise de Mathematique 5*, 111–13.

Tennant, N. (2000). Deductive versus expressive power: A pre-Gödelian predicament. *Journal of Philosophy 47*, 257–77.

van den Dries, L. (1988). Alfred Tarski's elimination theory for real closed fields. *Journal of Symbolic Logic 53*, 7–19.

van Heijenoort, J. (1967). Logic as calculus and logic as language. *Boston Studies in Philosophy of Science 3*, 440–46.

Veblen, O. (1902). Hilbert's foundations of geometry. *The Monist*, 303–9.

Veblen, O. (1904). A system of axioms for geometry. *Transactions of the American Mathematical Society 5*, 343–84.

Veblen, O. (1906). The foundations of geometry: An historical sketch and a simple example. *Popular Science Monthly 68*, 21–28.

Veblen, O. and J. W. Young (1908). A set of assumptions for projective geometry. *American Journal of Mathematics 30*, 347–80.

Veblen, O. and J. W. Young (1910). *Projective Geometry*, Volume 1. Boston: Ginn and Company.

Weyl, H. (1926). *Philosophie der Mathematik und Naturwissenschaften.* München: Oldenbourgh.

Whitehead, A. and B. Russell (1910). *Principia Mathematica*, Volume 1. Cambridge: Cambridge University Press.

Zach, R. (1999). Completeness before Post: Bernays, Hilbert, and the development of propositional logic. *Bulletin of Symbolic Logic 5*, 331–66.