

PDF version of the entry
Second-order and Higher-order Logic
<http://plato.stanford.edu/archives/fall2015/entries/logic-higher-order/>
from the FALL 2015 EDITION of the

STANFORD ENCYCLOPEDIA OF PHILOSOPHY



Edward N. Zalta	Uri Nodelman	Colin Allen	R. Lanier Anderson
Principal Editor	Senior Editor	Associate Editor	Faculty Sponsor

Editorial Board
<http://plato.stanford.edu/board.html>

Library of Congress Catalog Data
ISSN: 1095-5054

Notice: This PDF version was distributed by request to members of the Friends of the SEP Society and by courtesy to SEP content contributors. It is solely for their fair use. Unauthorized distribution is prohibited. To learn how to join the Friends of the SEP Society and obtain authorized PDF versions of SEP entries, please visit <https://leibniz.stanford.edu/friends/>.

Stanford Encyclopedia of Philosophy
Copyright © 2015 by the publisher
The Metaphysics Research Lab
Center for the Study of Language and Information
Stanford University, Stanford, CA 94305

Second-order and Higher-order Logic
Copyright © 2015 by the author
Herbert B. Enderton

All rights reserved.

Copyright policy: <https://leibniz.stanford.edu/friends/info/copyright/>

Second-order and Higher-order Logic

First published Thu Dec 20, 2007; substantive revision Wed Mar 4, 2009

Second-order logic is an extension of first-order logic where, in addition to quantifiers such as “for every object (in the universe of discourse),” one has quantifiers such as “for every *property* of objects (in the universe of discourse).” This augmentation of the language increases its expressive strength, without adding new non-logical symbols, such as new predicate symbols. For classical extensional logic (as in this entry), properties can be identified with sets, so that second-order logic provides us with the quantifier “for every *set* of objects.”

There are two approaches to the semantics of second-order logic. They differ on the interpretation of the phrase “for every set of objects.” Does this have some fixed meaning to which we can refer, or do we need to consider the variety of meanings the phrase might have? In the first case (which will be called standard semantics), we are taking for granted certain mathematical concepts. In the second case (which will be called general semantics), much less is being taken for granted. In this case, to be considered *valid*, a sentence will need to be true under all the allowable meanings of the phrase “for every set of objects.”

- 1. Syntax and translation
- 2. Standard semantics
- 3. General semantics
- 4. Higher-order logic
- 5. Systems of second-order number theory
- Bibliography
- Academic Tools
- Other Internet Resources
- Related Entries

1. Syntax and translation

In symbolic logic, the formula $(Px \rightarrow Px)$ will be true, no matter what object in the universe of discourse is assigned to the variable x . Which is to say nothing but that $\forall x(Px \rightarrow Px)$ will be true, no matter what subset of the universe of discourse is used to interpret the predicate symbol P . But is not that to say nothing but that the formula $\forall P \forall x(Px \rightarrow Px)$ is true, no matter what?

In first-order languages, there are some things we can say, and some that we cannot. Suppose, for example, that we want to express facts about the arithmetic of the natural numbers. That is, we want to express facts about the structure $(N; 0, S, <, +, \times)$ consisting of the set $N = \{0, 1, \dots\}$ of natural numbers, together with the common arithmetical operations and relations.

And we want to use a first-order language with quantifiers \forall interpreted as “for every natural number” and \exists interpreted as “for some natural number.” Moreover, we include in the language a constant symbol 0 for the number zero, a one-place function symbol S for the successor operation (which applied to a natural number gives the next one), a two place predicate symbol $<$ for the ordering relation $<$ on N , and two-place function symbols $+$ and \times for addition and multiplication, respectively.

With this language, we can now symbolize many of the facts we know to be true about the natural numbers. We can form the sentence $\forall x(x < Sx)$ expressing the fact that each number is smaller than the next one, for example. But a difficulty arises if we want to express the “well-ordering property” that any non-empty set of natural numbers has a smallest member. If P is a new one-place predicate symbol, then

$$\exists x Px \rightarrow \exists x(Px \ \& \ \forall y(Py \rightarrow (y = x \vee x < y)))$$

expresses the idea that P is true of some smallest number, if it is true of

any numbers at all. This formula is true in the structure $(N; 0, S, <, +, \times)$ when we interpret the predicate symbol P as being true of the numbers in some particular set—no matter what that set is—it says that the set has a least member, if it is non-empty. By now adding the quantifier $\forall P$

$$\forall P[\exists x Px \rightarrow \exists x(Px \ \& \ \forall y(Py \rightarrow (y = x \vee x < y)))]$$

we get a formalization of the well-ordering property.

The language of second-order logic extends the language of first-order logic by allowing quantification of predicate symbols and function symbols. As the foregoing example shows, in a second-order language for arithmetic, we can say that the natural numbers are well ordered. We know that the well-ordering property is not expressible by any first-order sentence, because the non-standard models of the (first-order) theory of $(N; 0, S, <, +, \times)$ are never well ordered. So going to second-order logic is a genuine extension. That is, we can translate some natural-language sentences, such as “The relation $<$ is a well-ordering,” into the language of second-order logic that are not translatable into the language of first-order logic.

For another example, we can (using choice) say that the universe of discourse is infinite by saying that there is a transitive relation on the universe such that every element bears the relation to something, but not to itself:

$$\exists R[\forall x \forall y \forall z(Rxy \ \& \ Ryz \rightarrow Rxz) \ \& \ \forall x[\neg Rxx \ \& \ \exists y Rxy]]$$

Here the second-order quantifier “ $\exists R$ ” expresses the existence of some binary relation on the universe. Because the only predicate symbol, R , is in the scope of this quantifier, the sentence has no predicate symbols open to interpretation. As is well known, no first-order sentence has for its models exactly the infinite structures.

In more detail, here is what is meant by a second-order language: One starts with a first-order language, and augments it by an unending supply of n -place predicate variables for each positive integer n , and an unending supply of n -place function variables for each positive integer n . (The function variables can be avoided, but that is another matter.) The formation rules for well-formed formulas are the obvious ones; in particular universal and existential quantification is allowed for any variable, be it an individual variable, a predicate variable, or a function variable.

To return to the example of natural numbers, we can express the Peano induction postulate by a second-order sentence:

$$\forall X[X0 \ \& \ \forall y(Xy \rightarrow XSy) \rightarrow \forall y Xy]$$

This sentence expresses the idea that X is true of all natural numbers, if it is true of 0 and its truth at some number y guarantees its truth at the successor of y , no matter what set of numbers X might be true of.

For an example involving the set \mathbf{R} of real numbers with its usual ordering, we can express the least-upper-bound property by a second-order sentence:

$$\forall X[\exists y \forall z(Xz \rightarrow z \leq y) \ \& \ \exists z Xz \rightarrow \exists y \forall y'(\forall z(Xz \rightarrow z \leq y') \leftrightarrow y \leq y')]$$

The foregoing examples are drawn from mathematical situations. There is also the intriguing possibility of *natural-language* sentences that seem to require second-order formulas for their formalization. George Boolos suggested the example, “There are some critics who admire only each other.” This sentence asserts the existence of a *set* of individuals having a certain property; it does not entail, for example, that there are *two* critics who admire each other and admire no others.

2. Standard semantics

Implicit in the previous section is the concept of *truth* of a second-order sentence in a structure. Consider a structure $M = (A, R, \dots)$ consisting of a non-empty set A serving as the universe of discourse, and some relations and functions on A interpreting the non-logical symbols. Then we want to count a second-order sentence of the form $\forall P \phi$ (where P is a k -place predicate variable) as being true in this structure if for *every* set Q of k -tuples of members of A , we have that ϕ is true in the structure when P is assigned the relation Q .

More formally, we need to define inductively what it means for a second-order formula ϕ to be *satisfied* in a structure $M = (A, R, \dots)$ under an assignment s of objects to the free variables in ϕ , which will be written $M \models \phi[s]$. The definition proceeds exactly as in the first-order case, except for the additional clauses for the second-order quantifiers. For a k -place predicate variable P ,

$$M \models \forall P \phi[s] \text{ iff for every } k\text{-ary relation } Q \text{ on } A, \text{ we have } M \models \phi[s']$$

where s' differs from s only in assigning the relation Q to the predicate variable P . (Here “iff” abbreviates “if and only if.”) Similarly, for a k -place function variable F ,

$$M \models \forall F \phi[s] \text{ iff for every } k\text{-place function } G \text{ on } A, \text{ we have } M \models \phi[s']$$

where s' differs from s only in assigning the function G to the function variable F . Observe that this definition refers, in the case of a 1-place predicate variable, to *all* subsets of A , that is, to the entire power set of A . It is this feature that accounts for the extraordinary semantical strength of second-order languages.

In the case of a second-order *sentence* σ (i.e., a formula with no free variable), the assignment s is no longer relevant, and we may speak unambiguously of the truth or falsity of σ in the structure M (that is, we can say that M is or is not a *model* of σ). In particular, the examples in §1 of translations from natural language into the language of second-order logic can now be seen to accomplish their intended purposes. The conjunction of the axioms for a linear ordering and the sentence

$$\exists x Px \rightarrow \exists x(Px \ \& \ \forall y(Py \rightarrow (y = x \vee x < y)))$$

is true in a structure $(A, <)$ iff the relation $<$ well-orders the set A . The sentence

$$\exists R[\forall x \forall y \forall z(Rxy \ \& \ Ryz \rightarrow Rxz) \ \& \ \forall x(\neg Rxx \ \& \ \exists y Rxy)]$$

is true in a structure iff the universe of discourse is an infinite set. This example shows that the compactness theorem does *not* hold for second-order logic. If we call the above sentence λ^∞ and we let λn be the first-order sentence “there are at least n different things in the universe,” then the set

$$\{\neg\lambda^\infty, \ \lambda 2, \ \lambda 3, \ \lambda 4, \ \dots\}$$

has no model, although each finite subset has a model.

The conjunction of the Peano postulates

$$\forall x(\neg 0 = Sx) \quad \text{and} \quad \forall x \forall y(Sx = Sy \rightarrow x = y)$$

and the Peano induction postulate

$$\forall X[X0 \ \& \ \forall y(Xy \rightarrow XSy) \rightarrow \forall y Xy]$$

is true in a structure (A, f, e) iff this structure is isomorphic to $(\mathbf{N}, S, 0)$, the natural numbers with the successor operation S and distinguished element 0. The conjunction of these three sentences provides an example of a sentence that is *categorical*, that is, it has exactly *one* model, up to isomorphism. By contrast, a first-order sentence can be categorical only if its one model is finite.

Similarly, the ordered field of real numbers, $(\mathbf{R}, 0, 1, +, \times, <)$, can be characterized up to isomorphism by the first-order axioms for an ordered field, together with the second-order sentence expressing the least-upper-bound property. It is well known that any model of these sentences must be isomorphic to the ordered field of real numbers. This example shows that the Löwenheim–Skolem theorem does *not* hold for second-order logic.

The preceding examples show that two everyday mathematical structures, $(\mathbf{N}, S, 0)$ and $(\mathbf{R}, 0, 1, +, \times, <)$ are *second-order characterizable*. That is, each one has a single second-order axiom of which it is the only model, up to isomorphism. One might ask what other structures might be second-order characterizable. Of course, there can be only countably many such structures, up to isomorphism, because each one needs a sentence.

Next, suppose that in these examples, we existentially quantify all the non-logical symbols (i.e., all the predicate symbols and function symbols). Where $\pi(0, S,)$ is the conjunction of the three Peano postulates, the sentence $\exists x \exists F \pi(x, F)$ is a sentence in the second-order language of equality, that is, it has no non-logical symbols at all.

A structure for the language of equality consists simply of a non-empty universe of discourse; there are no relations or functions or distinguished elements. Such a structure is of course determined up to isomorphism simply by its cardinality. For a sentence σ in the language of equality, let

its *spectrum* be the class of cardinalities in which it is true. For example, the spectrum of a valid sentence is the class of all non-zero cardinal numbers. The spectrum of an unsatisfiable sentence is empty. The spectrum of $\neg\sigma$ is the complement (relative to the class of all non-zero cardinal numbers) of the spectrum of σ . Conjunction and disjunction of sentences yield intersection and union of spectra. A sentence in the language of equality is determined up to logical equivalence by its spectrum.

The sentence $\exists x \exists F \pi(x, F)$ we made from the Peano postulates is true in the countably infinite cardinality and no other one. Thus its spectrum is a singleton. Say that a cardinal number κ is *second-order characterizable* if there is a sentence of the second-order language of equality that is true in cardinality κ and only there. (The non-zero finite cardinals are all first-order characterizable.) We have seen that the countable infinite cardinal is second-order characterizable. Similarly, we can show that the power of the continuum is second-order characterizable. Where $q(0, 1, +, \times, <)$ is the second-order sentence that characterizes the ordered field of real numbers up to isomorphism, the sentence

$$\exists x \exists y \exists F \exists G \exists R q(x, y, F, G, R)$$

is a sentence in the second-order language of equality that is true in the power of the continuum and in no other cardinality.

One might ask what other cardinal numbers are second-order characterizable. See the 1974 paper by S. Garland for an exploration of this question. There can be only countably many such cardinals, of course, because each one takes a sentence.

It is not hard to see that the least uncountable cardinal is second-order characterizable. We can use a sentence saying that the universe is infinite but not countable, and that any uncountable subset is equinumerous to the

entire universe. Thus we have sentences κ and λ in the second-order language of equality that characterize the least uncountable cardinal and the power of the continuum, respectively. The sentence $\kappa \leftrightarrow \lambda$ is a valid sentence if the continuum hypothesis is true, and only then. We can conclude that not all issues involving second-order logic are necessarily settled in ZFC.

The much-studied theory PA, first-order Peano arithmetic, is of course obtained by using, in place of the second-order induction postulate $\forall X[X0 \ \& \ \forall y(Xy \rightarrow XSy) \rightarrow \forall y Xy]$, the corresponding first-order *schema*

$$\phi(0) \ \& \ \forall y(\phi(y \rightarrow \phi(Sy)) \rightarrow \forall y \phi(y)$$

where ϕ can be any suitable first-order formula. The effect of this schema is well known; it assures that any *definable* set that contains 0 and is closed under successor must contain everything.

Suppose that, by analogy, we start from our second-order axiomatization of the ordered field of real numbers, and replace the second-order least-upper-bound axiom by the corresponding schema. The result is an infinite set of first-order axioms, assuring that any definable set that is non-empty and bounded has a least upper bound. The models of this are called *real-closed ordered fields*. Interestingly, this concept was first formulated by algebraists, not by logicians.

One measurement of the strength of second-order logic is the complexity of its set of valid sentences. Let V^1 be the set of valid sentences of first-order logic and let V^2 be the set of valid sentences of second-order logic. More specifically, in first-order logic with only a single 2-place predicate symbol P , we know that the set $V^1(P)$ of valid sentences is a complete computably enumerable set (i.e., a complete recursively enumerable set). (Here we can assign Gödel numbers and view $V^1(P)$ as a set of natural numbers, or equivalently we can view it directly as a set of words over a

finite alphabet.) And Tarski has pointed out that the set $V^1(=)$ of valid sentences in the first-order language of equality (with no non-logical symbols at all) is decidable.

For comparison, let $V^2(=)$ be the set of valid sentences in the second-order language of equality. What is the complexity of this set?

Let π be the conjunction of the Peano postulates and the recursion equations for addition and multiplication. Thus π is a second-order sentence in the language of arithmetic, with $0, S, +$, and \times . The sentence π is categorical; its only model is $(\mathbf{N}, 0, S, +, \times)$, up to isomorphism. Consequently, for a sentence σ in the language of arithmetic, σ is true in arithmetic iff the conditional $(\pi \rightarrow \sigma)$ is valid. This shows that $V^2(0, S, +, \times)$ cannot be arithmetical (i.e., cannot be first-order definable in arithmetic), lest truth in arithmetic be definable, in violation of Tarski's theorem. Now we can quantify away all the non-logical symbols; a sentence $\phi(P)$ is valid iff the sentence $\forall P \phi(P)$ is valid. The conclusion is that $V^2(=)$ is not arithmetical.

As interesting as that may be, it is merely the tip of the iceberg. To begin with, we can show that $V^2(=)$ is not analytical, that is, is not definable in arithmetic by a *second-order* formula. The proof of Tarski's theorem, showing that the set of true first-order sentences of arithmetic is not first-order definable in arithmetic, also shows that the set of true second-order sentences of arithmetic is not second-order definable in arithmetic. The rest of the argument is unchanged. And later we will see that even more is true.

In a very different direction, R. Fagin has shown a surprising connection between a topic in computational complexity and second-order definability over *finite* structures. For example, a finite graph can be regarded as a pair (V, E) consisting of a non-empty *vertex set* V and a symmetric *edge*

relation E on V . The statement that the graph can be properly colored with three colors can be expressed by a second-order sentence: there exist subsets R, G, B that partition V in such a way that two vertices connected by an edge are never the same color. This sentence is Σ -1-1, that is, it has the form

[existential second-order quantifiers] [first-order formula].

It is well known that being three-colorable is an NP property of a graph. That is, it is a property that is recognizable in polynomial time by a non-deterministic Turing machine. (There is a non-deterministic Turing machine M and a polynomial p such that whenever (V, E) , suitably encoded, is given to M , then if (V, E) is three-colorable then some computation of M will accept the graph within $p(n)$ steps, where n measures the size of (V, E) , and if (V, E) is not three-colorable then no computation of M will ever accept the graph.)

Fagin showed that this is not an isolated example; *every* NP property of finite graphs is definable by a Σ -1-1 sentence of second-order logic. And conversely, any Σ -1-1 sentence defines an NP property. And in place of graphs, we can use directed graphs or other finite structures. Fagin's theorem states that a property of finite structures is an NP property if and only if it is definable by a Σ -1-1 second-order sentence.

3. General semantics

A key feature of the “standard semantics” discussed in the previous section is that, for a one-place predicate variable X , the quantifier $\forall X$ ranges over the entire power set of the universe of discourse. We have seen that this feature gives second-order languages a high degree of expressive strength.

But do we really want the quantifier $\forall X$ to range over the actual power set? The predicativist will object that the power-set operation is not meaningful. And even the classical mathematician will admit that there are some obscure features of the power-set operation. The independence of the continuum hypothesis illustrates one such obscurity. If our goal is to study the foundations of mathematics, then it might be prudent not to take for granted that we already know all about power sets.

The concept of *general semantics* for second-order logic avoids any pretense that the power-set operation is a fixed well-understood resource. Instead, the range of the quantifier $\forall X$ must be directly specified.

By a *general pre-structure* for a second-order language we mean a structure in the usual sense (a universe of discourse plus interpretations for the non-logical symbols) together with the additional sets:

- The *n*-place *relation universe* for each positive integer *n*. This must be a collection of *n*-ary relations on the universe of discourse. In particular, the 1-place relation universe must be some collection of subsets of the universe. Thus it is part (perhaps all, perhaps not) of the power set of the universe.
- The *n*-place *function universe* for each positive integer *n*. This must be a collection of *n*-place functions on the universe of discourse.

For a general pre-structure *M*, there is a natural way to define what it means for a second-order formula ϕ to be *satisfied* in a structure *M* under an assignment *s* of objects to the free variables in ϕ , which again will be written $M \models \phi[s]$. The second-order quantifiers are now defined to range over the corresponding universe. For a *k*-place predicate variable *P*,

$$M \models \forall P \phi[s] \text{ iff for every } k\text{-ary relation } Q \text{ in the } k\text{-place relation universe, we have } M \models \phi[s']$$

where *s'* differs from *s* only in assigning the relation *Q* to the predicate variable *P*. Similarly, for a *k*-place function variable *F*,

$$M \models \forall F \phi[s] \text{ iff for every } k\text{-place function } G \text{ in the } k\text{-place function universe, we have } M \models \phi[s']$$

where *s'* differs from *s* only in assigning the function *G* to the function variable *F*. In the case of a second-order *sentence* σ (i.e., a formula with no free variable), the assignment *s* is no longer relevant, and we may speak unambiguously of the truth or falsity of σ in the general pre-structure *M* (that is, we can say that *M* is or is not a *model* of σ).

But for second-order logic, we do not really want the 1-place relation universe to be an *arbitrary* collection of subsets of the universe. We might not know everything about the power-set operation, but we know *some* things about it. In effect, using general pre-structures amounts to treating a second-order language as a many-sorted *first-order* language.

There are some subsets of the universe that we know about, because we can define them. That is, suppose ϕ is a formula in which only the variable *u* occurs free. Then the set ϕ *defines* in *M* is the set consisting of all members *a* of *M* such that ϕ is satisfied in *M* when *a* is assigned to *u*. This idea can be extended. Suppose that ϕ has only the free variables *u*, *v*, *w*, *x*, *Y*, and *Z* (where *Y* is an *m*-place predicate variable and *Z* is an *n*-place function variable). Suppose that *c* and *d* are members of the (individual) universe *IM* of *M*, that *E* is in *M*'s *m*-place relation universe, and that *F* is in *M*'s *n*-place function universe. Then the binary relation ϕ *defines* in *M* from the parameters *c*, *d*, *E*, and *F* is the set of pairs $\langle a, b \rangle$ of elements of *IM* such that ϕ is satisfied in *M* when its variables *u*, *v*, *w*, *x*, *Y*, and *Z* are assigned *a*, *b*, *c*, *d*, *E*, and *F*, respectively. That is, it is the binary relation:

$$\{\langle a, b \rangle \mid M \models \phi(u, v, w, x, Y, Z) [a, b, c, d, E, F]\}$$

Obviously, this concept can be generalized to the situation where a k -ary relation is defined from any particular number of parameters.

Then it is reasonable to restrict attention to general pre-structures that are closed under definability. Thus in the situation just described, it is reasonable to expect M 's 2-ary relation universe to contain the binary relation that ϕ defines from parameters in the pre-structure. That is, we expect the sentence

$$\forall w \forall x \forall Y \forall Z \exists R \forall u \forall v [Ruv \leftrightarrow \phi(u, v, w, x, Y, Z)]$$

to be true in M . Call such sentences *comprehension axioms*.

By a *general structure* for a second-order language (also called a *Henkin structure*) is meant a general pre-structure in which all comprehension axioms (for all formulas) are true. (Here ϕ might contain quantifiers over predicate variables, so that even impredicative comprehension axioms are to be true. In Section 5, we consider alternatives to “full comprehension.”) Among the general structures are those in which the 1-place relation universe is the actual power set of the individual universe, and so forth. (Call such a general structure *absolute*.) But there can be others.

We obtain the *general semantics* (also called the *Henkin semantics*) for a second-order language by considering all general structures. That is, for a sentence σ to be valid in the general semantics, it must be true in all general structures. This is a stronger requirement than saying that σ is valid in the standard semantics. A sentence that is valid in the standard semantics is true in those general structures for which the 1-place relation universe is the full power set of the individual universe, and so forth. But such a sentence σ might turn out to be false in some general structure (that is, $\neg\sigma$ might have a general model).

The main feature of the general semantics is a result of the “nothing but” type: Second-order logic with the general semantics is nothing but first-order logic (many-sorted) together with the comprehension axioms. Thus a sentence is valid in the general semantics iff it is logically implied (in first-order logic) by the set of comprehension axioms.

This reduction to first-order logic yields at once the following results:

- (Enumerability) In a second-order language with finitely many non-logical symbols, the set of sentences that are valid in the general semantics is computably enumerable. This holds because the set of comprehension axioms is a computable set (i.e., a recursive set).
- (Compactness) A set of sentences has a general model if every finite subset has a general model.
- (Löwenheim–Skolem) If a set of sentences has a general model, then it has a countable general model.

In each of these three cases, there is a sharp contrast to the situation of standard semantics considered in Section 2. Moreover, a deductive calculus can be given for second-order logic (adapted from first-order logic and augmented by the comprehension axioms) that will be complete for the general semantics.

For comparison, axiomatic set theory (ZFC say) is a *first-order* theory; a model of set theory must supply a power-set operation. But the language of set theory has certain higher-order aspects, in that it permits us to speak of sets, sets of sets, and so forth.

In the extreme case of a structure M in which *all* relations are definable (e.g. a structure with a one-point universe), the general semantics will coincide with the standard semantics.

4. Higher-order logic

There is no need to stop at second-order logic; one can keep going. We can add to the language “super-predicate” symbols, which take as arguments both individual symbols (either variables or constants) and predicate symbols. And then we can allow quantification over super-predicate symbols. And then we can keep going further.

(The reader is to be cautioned that there are in the literature two different ways of counting the order. According to one scheme, third-order logic allows super-predicate symbols to occur free, and fourth-order logic allows them to be quantified. According to the other scheme, third-order logic already allows quantification of super-predicate symbols.)

We reach the level of *type theory* after ω steps. And continuation into the transfinite is conceivable.

We have seen that, although the set V^1 of valid formulas of first-order logic is computably enumerable, the corresponding set V^2 for second-order logic (with the standard semantics) is vastly more complex. This phenomenon does not continue into the higher orders.

There is a sense in which the power-set operation is definable in second-order logic. Consider a language with a one-place predicate symbol I (for individuals), a one-place predicate symbol S (for sets), and a two-place predicate symbol E (for the membership relation). Then to express the idea that S is the power set of I we can use the conjunction (call it σ) of the following four sentences:

- $\forall x(Ix \rightarrow Sx)$ where “ \rightarrow ” denotes exclusive disjunction
- $\forall x \forall y (Exy \rightarrow Ix \& Sy)$
- $\forall x \forall y (Sx \& Sy \& \forall t(Ex \leftrightarrow Ety) \rightarrow x = y)$ (extensionality)
- $\forall X \exists y(Sy \& \forall t(Ix \rightarrow (Ex \leftrightarrow Xt)))$ (comprehension).

Clearly σ is true in any structure whose universe is the disjoint union of a set A and its power set $P(A)$ and which assigns A to I , assigns $P(A)$ to S , and assigns the membership relation \in to E . Conversely, let M be any model of σ (in the standard semantics). Let f be a one-to-one function from M 's interpretation of I onto a set A that is disjoint from its power set (there always is such a set). Extend f to all of M 's universe by defining for each s in M 's interpretation of S

$$f(s) = \{f(i) \mid M \models E[i, s]\}$$

(that is, $f(s)$ is the set of things in A that M thinks belong to s). Then f is an isomorphism from M to a structure whose universe is the disjoint union of a set A and its power set $P(A)$ and which assigns A to I , assigns $P(A)$ to S , and assigns the membership relation \in to E . So roughly speaking, σ defines the power-set operation “to within isomorphism.”

In a similar vein, we can define to within isomorphism the power set of $I \times I$, i.e., the set of binary relations on I . And so forth.

This second-order expressibility of the power-set operation permits the simulation of higher-order logic within second order. More specifically, we have the following result of Hintikka (1955): For each formula ϕ of higher-order logic (in a language with finitely many non-logical symbols), we can effectively find a sentence ψ of second-order logic (in the language of equality) such that ϕ is valid if and only if ψ is valid. The sentence ψ is constructed by first expanding the language by adding symbols for universes of various types (individuals, sets of individuals, ...) and for membership in these universes. Then ϕ 's validity is equivalent to the validity of a second-order formula

The universes are correctly arranged $\rightarrow \phi^*$

where ϕ^* is a suitable relativization of ϕ . Finally, we can prefix universal quantifiers to obtain a sentence ψ in the language of equality.

Thus the set of validities of seventeenth-order logic is computably reducible to $V^2(=)$, the set of second-order validities in the language of equality. (In fact, these two sets are computably isomorphic.) So in this aspect, the complexity of higher-order logic does not increase with the order. It follows that the set $V^2(=)$ has a high degree of complexity. We can extend our earlier observation that it is not definable in second-order arithmetic; it is not definable in arithmetic of higher order either. Montague in 1965 extended this into the transfinite. At the time, he was heard to say that the set $V^2(=)$ does not lie in any Kleene hierarchy, “past, present, or future.”

The fact that we can express the power-set operation in second-order logic (and can iterate the procedure) gives second-order logic some large part of the expressiveness of set theory. Quine has claimed that second-order logic is not really logic, but rather set theory in disguise. And Robert Vaught has commented that studying second-order logic was like studying “the standard model of set theory.”

The complexity of $V^2(=)$ does not change much if we impose limitations on the quantifier form of the sentences. The foregoing reduction of higher-order logic yields Π -1-2 sentences, so we can conclude that the set of valid Π -1-2 sentences in the language of equality is computably isomorphic to the full $V^2(=)$. That is about the best possible result. The set of Π -1-1 valid sentences is a complete computably enumerable set; once we drop the universal second-order quantifiers we are looking at first-order formulas. The set of Σ -1-1 valid sentences in the language of equality is a complete co-c.e. set, i.e., the complement of a computably enumerable set.

(The sentence $\exists P\phi(P)$ is true in every non-zero cardinality iff the elementary sentence $\phi(P)$ has models of every finite size, a co-c.e.

property. And to a Turing machine we can effectively assign an elementary sentence having models of every finite size iff the machine never halts.) The set of valid Σ -1-2 sentences in the language of equality is also computably isomorphic to the full $V^2(=)$, but this fact requires a different sort of proof.

5. Systems of second-order number theory

The language of arithmetic (with 0, S , $<$, $+$, and \times) is important to the foundations of mathematics. As axioms, we can take the usual Peano postulates, including the second-order induction axiom. In the standard semantics, the only model of the Peano postulates, up to isomorphism, is the usual model of arithmetic. So the theory generated by these axioms (in the standard semantics) is simply the second-order theory of true arithmetic.

But suppose we consider instead the *general* semantics. The Peano postulates have general models that can differ from the usual model in either (or both) of two ways. We can employ the compactness theorem to construct non-standard general models of the Peano postulates containing infinitely large numbers. We can also find general models of the Peano postulates in which the universe of sets is less than the full power set of the individual universe (i.e., general models that are not absolute). Indeed, any countable general model must be of this kind.

In the context of general models, we add the additional axiom schema for *choice*:

$$\forall n \exists X \phi(n, X) \rightarrow \exists Y \forall n \phi(n, \{t \mid Ynt\})$$

prefixed by universal quantifiers as needed. (Here the formula that has been written as $\phi(n, \{t \mid Ynt\})$ is obtained from $\phi(n, X)$ by replacing each term Xu by the term Ynu .)

The Peano postulates are strong enough to provide us with pairing functions. Consequently, for a general model, its 1-place relation universe completely determines its k -place relation universe for each k .

The traditional terminology is to refer to second-order number theory as *analysis*. The name derives from the fact that it is possible to identify real numbers with sets of natural numbers. The second-order quantifiers over sets of natural numbers can then be viewed as quantifiers over the real numbers. The appropriateness of the name is open to question, but its usage is well established. Accordingly, by a *model of analysis* we will mean a general model of the Peano postulates with choice. As a general model, any model of analysis must of course satisfy all of the comprehension axioms; later we will consider weakening this requirement.

Let A_2 be the theory generated by the Peano postulates with choice (in the general semantics). Then A_2 is a complete computably enumerable subset of true second-order arithmetic. A_2 contains every true Σ_0^1 sentence. It does not contain every true Π_0^1 sentence.

We can obtain a stronger theory by restricting attention to the models of analysis that differ from the usual model in only the second of the two ways described previously. That is, define an ω -*model of analysis* to be a model of analysis in which the individual universe is the actual set of natural numbers and the symbols 0 and S have their usual interpretations. (Consequently, the symbols $<$, $+$, and \times have their usual interpretations.) The set universe of an ω -model of analysis must therefore be some part (possibly all) of the power set of the natural numbers, but it must be such that full comprehension is satisfied.

The motivation for considering ω -models can be stated as follow. We have a clear understanding—or so we like to think—of the set of natural

numbers. But we do not have anything like the same understanding of the power set of the set of natural numbers. So it is reasonable to hold fixed the part we are sure of, but to leave open to interpretation the part we are not so sure about.

An ω -model of analysis is completely determined by its set universe. Because we have pairing functions that are first-order definable, the set universe will determine the binary relation universe, and so on. A Löwenheim–Skolem argument will show that there are ω -models of analysis with a countable set universe.

In any ω -model of analysis, the true *first-order* sentences are exactly the ones true in usual model. But ω -models can differ on second-order sentences. Let A_ω be the theory of ω -models, that is, the set of sentences true in all ω -models of analysis. Then A_ω extends A_2 . It is a complete Π_1^1 set. It contains every true Π_1^1 sentence; it does not contain every true Σ_1^1 sentence.

By the ω -rule we mean the infinitary rule of inference that infers $\forall x \phi(x)$ from the premisses $\phi(0), \phi(1), \phi(2), \dots$. Suppose we add this rule to the usual logical apparatus, and see what is then deducible from the Peano postulates with choice. A “deduction” using the ω -rule will in general be infinite, but it must be well founded. It is not hard to see that any sentence deducible in this way is in A_ω . Conversely, a completeness theorem holds: Every sentence in A_ω has a deduction of the sort described.

In a 1961 paper, Andrzej Mostowski described a way of going one step further, toward a stronger theory still. Suppose we are willing to consider not just the order type ω as being sufficiently understood, but even the concept of being a well-ordering. Define a β -model of analysis to be an ω -model with the additional property that orderings in the model that

appear to be well-orderings really are. That is, an ω -model M of analysis is a β -model if every ordering relation (on the natural numbers) in M with the property that non-empty sets in M always have least members, is in fact a well-ordering.

Then the set $A\beta$ of sentences that are true in all β -models turns out to be a complete Π -1-2 set. It contains every true Π -1-2 sentence; it does not contain every true Σ -1-2 sentence. Clearly we have the inclusions

$$A2 \subseteq A\omega \subseteq A\beta \subseteq \text{True Second-Order Arithmetic.}$$

For example, the number-theoretic part of a transitive \in -model of ZF set theory will always be a β -model of analysis. But we can construct a β -model without strong assumptions. There is in fact a *smallest* β -model, and it can be constructed in a way reminiscent of Gödel's definition of the class L of constructible sets.

Finally, it should be mentioned that there are contexts in which theories of second-order arithmetic with *less* than full comprehension are suitable. For example, one can take the theory ACA given by the Peano postulates together with comprehension axioms for *first-order* formulas only. Such theories have been shown to be applicable to the study of “reverse mathematics.”

Bibliography





- Boolos, George, 1975. “On Second-Order Logic,” *The Journal of Philosophy*, 72: 509–527. Also in *Logic, Logic, and Logic*, by George Boolos, Cambridge, MA: Harvard University Press, 1998, pp. 37–53.
- Church, Alonzo, 1956. *Introduction to Mathematical Logic*, Volume 1. Princeton: Princeton University Press.
- , 1972. “Axioms for Functional Calculi of Higher Order,” in *Logic*

and Art: *Essays in Honor of Nelson Goodman*, Richard S. Rudner and Israel Scheffler (eds.), Indianapolis and New York: Bobbs-Merrill Company, pp. 197–213.

- Enderton, Herbert B., 2001. *A Mathematical Introduction to Logic*, second edition. San Diego: Academic Press.
- Fagin, Ronald, 1974. “Generalized First-Order Spectra and Polynomial-Time Recognizable Sets,” in *Complexity of Computation*, Richard M. Karp (ed.), SIAM-AMS Proceedings, vol. 7, Providence: American Mathematical Society, pp. 43–73.
- Garland Stephen J., 1974. “Second-Order Cardinal Characterizability,” in *Axiomatic Set Theory*, Thomas J. Jech (ed.), Proceedings of Symposia in Pure Mathematics, vol. 13, part 2, Providence: American Mathematical Society, pp. 127–146.
- Grzegorzczak, Andrzej, A. Mostowski, and C. Ryll-Nardzewski, 1958. “The Classical and the ω -Complete Arithmetic,” *The Journal of Symbolic Logic*, 23: 188–205.
- Henkin, Leon, 1950. “Completeness in the Theory of Types,” *The Journal of Symbolic Logic*, 15: 81–91.
- Hintikka, K. Jaakko, 1955. “Reductions in the Theory of Types,” in *Two Papers on Symbolic Logic*, Acta Philosophica Fennica, No. 8, Helsinki, pp. 57–115.
- Montague, Richard, 1965. “Reductions of Higher-Order Logic,” in *The Theory of Models*, J. W. Addison, Leon Henkin, and Alfred Tarski (eds.), Amsterdam: North-Holland Publishing Co., pp. 251–264.
- Mostowski, Andrzej, 1961. “Formal System of Analysis Based on an Infinitistic Rule of Proof,” in *Infinitistic Methods*, Warsaw: Państwowe Wydawnictwo Naukowe, and Oxford, London, New York, and Paris: Pergamon Press, pp. 141–166.
- Orey, Steven, 1956. “On ω -Consistency and Related Properties,” *The Journal of Symbolic Logic*, 21: 246–252.
- Shapiro, Stewart, 1991. *Foundations without Foundationalism: A Case*

- for Second-Order Logic*. Oxford: Oxford University Press.
- Simpson, Stephen, 1999. *Subsystems of Second Order Arithmetic*. Berlin: Springer.
- Väänänen, Jouko, 2001. “Second-Order Logic and Foundations of Mathematics,” *The Bulletin of Symbolic Logic*, 7: 504–520.

Academic Tools

-  How to cite this entry.
-  Preview the PDF version of this entry at the Friends of the SEP Society.
-  Look up this entry topic at the Indiana Philosophy Ontology Project (InPhO).
-  Enhanced bibliography for this entry at PhilPapers, with links to its database.

Other Internet Resources

[Please contact the author with suggestions.]

Related Entries

computability and complexity | plural quantification | type theory

Copyright © 2015 by the author
Herbert B. Enderton