Second-Order Logic

Péter Mekis

May 17, 2016

Contents

1 Introduction

Consider the following argument:

(1) Jumbo is an elephant.
Elephant is a species.
Therefore: Jumbo is a species.

This argument is not valid, since we can safely assume that both of its premises are true, while assuming that its conclusion is not. Ordinary intuition suggests that the reason is that there is an incompatibility between the property *species* and the individual *Jumbo* in the conclusion. However, first-order analysis doesn't support this intuition. Jumbo is an individual, while being an elephant and being a species are properties; so we can translate them by the individual constant a, and the predicate symbols P and Q, respectively. The result is

(1a) $P(a)$
$Q(P)$
$\therefore Q(a)$

Here, something is obviously ill-formed; but rather than the conclusion, it is the the second premise. In the syntax of first-order predicate logic only individual terms are allowed as arguments of predicate symbols. The mistake in the analysis is easy to spot; being a species is not a property of individuals, like Jumbo, but a property of properties, like elephant. We call such properties *second-order properties*; and express them with *second-order predicate symbols*. Using calligraphic letters is a common notational convention for these. Second-order predicate symbols take ordinary, first-order predicate symbols as their arguments; but not individual terms. If we use a second-order predicate in our analysis,

(1b) $P(a)$
$\mathcal{Q}(P)$
$\therefore \mathcal{Q}(a)$

—it will be obvious that the problematic part of the argument is its conclusion, since it contains a second-order predicate with an individual term in its argument position.

Let us see a second example now.

(2) Santa Claus has all characteristic properties of a pedophile.

First of all, let us stress that we don't claim that Santa would be a pedophile. By characteristic properties we mean properties that people usually use to identify members of a group; but those criteria might be deceptive. So, using the unary first-order predicate P for pedophile, and a for Santa, the following is neither the proper translation of the original sentence, nor a consequence of it:

(2a) Pa

In fact, the original sentence contains a generalization over properties. Since properties are expressed by either open formulas or unary predicates, our translation needs to be either a first-order formula scheme, like the induction axiom of Peano Arithmetic, or one that contains bound predicate variables. The latter version is called *second-order quantification*. Also, we can observe that being a characteristic property is a relation between properties, one being a characteristic of the other. Just as in the previous example, it amounts to using a second-order predicate symbol in the translation. This prevents us from expressing ourselves in a first-order scheme; so we have to use second-order quantification, too. Let us use X as a unary predicate variable, a and P as in (2a), and $\mathcal{R}$ for being a characteristic property of another property. With these symbols, our translation will take the following form:

(2b) $\forall X\ \big(\mathcal{R}(X,P) \to P(a)\big)$

This is a full-blooded second-order formula that exhibits all the syntactical features that we are going to use in what follows.

2 Syntax

A second-order language is an extension of a first-order language with the following logical and extralogical tools:

- We expand the signatures of our languages by second-order predicates. They express properties and relations of properties and relations (of individuals). We use calligraphic letters to distinguish them from first-order predicates: $\mathcal{P}, \mathcal{Q}, \mathcal{R}, \mathcal{S}$. Note that besides the arity of the predicate the arities of the arguments also matter. These arities can be indicated in subscripts: $\mathcal{R}_{1,1}$, $\mathcal{R}_{2,2}$, and $\mathcal{R}_{1,2}$ stand for a second-order binary relation of first-order properties, one of first-order binary relations, and one of a first-order property and a first-order binary relation, respectively. We omit the subscripts whenever they are obvious from the context.

- We introduce predicate variables. Their values are extensions of predicates. We use uppercase letters to distinguish them from individual variables: X, Y, Z. Their arities can be indicated in subscripts, too; X_1 is unary, Y_2 is binary.

- besides the usual atomic formulas we have the following new ones. Let $\mathcal{R}_{n_1,\dots,n_k}$ be a second-order predicate symbol, and let X_n be a first-order predicate variable. Let $T_1, \dots, T_k$ be predicate symbols or predicate variables (we can call them *predicate terms*), and let $t_1, \dots, t_n$ be individual terms. Then
$$\mathcal{R}_{n_1,\dots,n_k}(T_1, \dots, T_k)$$
and
$$X_n(t_1, \dots, t_n)$$
are atomic formulas.

- Second-order quantifiers bind predicate variables, and they range over extensions of predicates. We use the same symbols as for first-order quantifiers; it is the uppercase variables that make the difference: second-order quantified formulas take the forms $\forall X\ \varphi$, and $\exists Y\ \varphi$.

- Besides the above tools, second-order languages might contain second-order function symbols with ordinary first-order functions as their arguments, function variables, and quantification over functions. The arguments of a second-order function symbol might even include first-order predicate symbols; and the arguments of a second-order predicate symbol might even include first-order function symbols. We don't introduce these tools in the present discussion.

3 Standard semantics

- To get a second-order structure, we need to expand a first-order structure with the interpretations of second-order predicate symbols. Let R be a k-ary second-order predicate symbol with n_i-ary argument at the ith argument place. Its ith argument's extension is a subset of A^{n_i}; and $\mathcal{R}$'s extension consists of k-tuples of such extensions. Formally:
$$\mathcal{R}^{\mathcal{A}}_{n_1,\dots,n_k} \subseteq \mathcal{P}(A^{n_1}) \times \dots \times \mathcal{P}(A^{n_k})$$
(Here $\mathcal{P}(A)$ is the power set of A. Note the clashing notations; we use calligraphic $\mathcal{P}$ for power set in the metalanguage, while the same symbol

would stand for a second-order predicate symbol in the object language. To avoid confusion, we don't use $\mathcal{P}$ for second-order predicates in this section.)

- The value of an n-ary predicate variable is an n-ary predicate extension; that is, a set of ordered n-tuples of individuals:

$$g(X_n) \subseteq A^n$$

- Let us define the value of a predicate term T as

$$|T|_g^{\mathcal{A}} = \begin{cases} T^{\mathcal{A}} \text{ if } T \text{ is a predicate symbol;} \\ g(T) \text{ if } T \text{ is a predicate variable.} \end{cases}$$

Now the truth conditions of the new atomic formulas are:

$$\mathcal{A} \models_g \mathcal{R}(T_1, \ldots, T_k) \quad \Leftrightarrow \quad \langle |T_1|_g^{\mathcal{A}} \ldots |T_k|_g^{\mathcal{A}} \rangle \in \mathcal{R}^{\mathcal{A}}$$

$$\mathcal{A} \models_g \mathcal{X}(t_1, \ldots, t_n) \quad \Leftrightarrow \quad \langle |t_1|_g^{\mathcal{A}} \ldots |t_n|_g^{\mathcal{A}} \rangle \in g(X)$$

- Finally the truth conditions of the second-order-quantified formulas:

$$\mathcal{A} \models_g \forall X \ \varphi \quad \Leftrightarrow \quad \mathcal{A} \models_{g'} \varphi \text{ for every } g' \text{ such that } g'[x]g$$

$$\mathcal{A} \models_g \exists X \ \varphi \quad \Leftrightarrow \quad \mathcal{A} \models_{g'} \varphi \text{ for at least one } g' \text{ such that } g'[x]g$$

Truth without respect to assignment, as well as logical consequence, are defined as usual:

$$\mathcal{A} \models \varphi \quad \Leftrightarrow \quad \text{for every } g: \ \mathcal{A} \models_g \varphi$$

$$\Gamma \models \varphi \quad \Leftrightarrow \quad \text{for every } \mathcal{A}: \ \text{if for every } \psi \in \Gamma \mathcal{A} \models \psi, \text{ then } \mathcal{A} \models \varphi$$

The theory of second-order languages with the standard semantics and a corresponding definition of logical consequence relation is called standard second-order logic, or simply second-order logic, dhortened as SOL.

Besides the standard semantics, there are several weaker semantics for standard second-order languages; most notably the Henkin semantics. We don't discuss them here.

4 Comprehension

A very important new logical law of SOL is the law of *comprehension.* Whatever the formula φ is,

$$\exists Y \ \forall x \ \big(Y(x) \leftrightarrow \varphi(x)\big)$$

The schema asserts that whatever property is definable by a formula in a second-order language, its extension is a possible value of a unary predicate variable. The law is a consequence of the separation schema of the set theory that we use as a semantic metatheory; every formula of the object language defines a subset of the domain.

Note that the converse of the law is not true: it is not the case that for every subset of the domain, there is a formula of the object language that defines it. This asymmetry is the reason why SOL is essentially richer than FOL; a lot more can be expressed by second-order quantified formulas than by first-order schemes.

5 Nonfirstorderizability

In what follows we will compare the expressive powers of a first-order language and the corresponding second-order language with the same signature. For the sake of comparability, we don't use second-order predicates; only predicate variables and second-order quantification.

In some cases it is possible to find a first-order equivalent for a second-order formula. That is, given a formula ϕ that contains second-order quantification, sometimes there is formula without second-order quantification which is true in the very same structures; but in a great many cases it is not possible. Most of the well-known examples for such formulas are technical; but there are some from ordinary language, too. Let us first see one of these.

The following is known as the Geach-Kaplan sentence:

(GK) Some critics admire only one another.

(GKB) $\exists Z\ \big(\exists x\ Zx \wedge \forall x\ \forall y\ ((Zx \wedge xRy) \rightarrow (Zy \wedge x \neq y))\big)$

This formula has no first-order equivalent. This can be shown by translating it into the language of Peano Arithmetic, by the following definition:

$$xRy \ \overset{\text{def}}{\Leftrightarrow}\ x = 0 \vee x = sy$$

Now the formula reads as

(GKP) $\exists Z\ \big(\exists x\ Zx \wedge \forall x\ \forall y\ ((Zx \wedge (x = 0 \vee x = sy)) \rightarrow (Zy \wedge x \neq y))\big)$

Let us see the truth conditions of this sentence. It claims that there is a nonempty set Z of natural numbers such that they satisfy the condition

$$\forall x\ \forall y\ ((Zx \wedge (x = 0 \vee x = sy)) \rightarrow (Zy \wedge x \neq y))$$

Now can 0 be an element of Z? Assume it is. Then we have

$$\forall y\ ((Z0 \wedge (0 = 0 \vee 0 = sy)) \rightarrow (Zy \wedge 0 \neq y))$$

The antecedent of the conditional is true partly by assumption and partly by tautology, so we can eliminate it and we get the simpler formula

$$\forall y\ (Zy \wedge 0 \neq y))$$

This, on its turn, has the instance

$$Z0 \wedge 0 \neq 0)$$

—which is a logical contradiction.

So, Z is nonempty, and 0 is not in Z, so Z has at least one nonzero member. Let us pick one; say, a_0. We proved in PA before that every nonzero number has a predecessor. According to the formula, this predecessor is in Z, too; so that must be nonzero, too. Let it be a_1. Iterating this reasoning, we get that there is an infinite sequence of numbers $a_0, a_1, a_2, \ldots$ such that $a_0 = sa_1$, $a_1 = sa_2$, and so on.

Thus the formula claims that Z contains an infinite descending chain of numbers; or, in other terms, it claims that the set of natural numbers is not well-founddded with respect to the < relation as we defined it in PA. We saw discussing PA that neither the existence nor the nonexistence of such a chain can be expressed in first-order logic; but as it turns out, it can in second-order logic.

There are of course more direct ways to make this claim. By definition the well-foundedness of a set Z with respect to a relation R means that every nonempty subset Y of Z has a minimal element with respect to R:

$$\mathcal{W}(Z) \overset{\text{def}}{\Leftrightarrow} \forall Y \ \big((\forall x \ (Yx \to Zx) \land \exists x \ Yx\big) \to \exists x \ \big(Yx \land \forall x' \ (Yx' \to \neg\, x'Rx)\big)\big)$$

Yet another way to say that there aren't nonstandard numbers is to claim that every number is within the intersection of all sets of numbers that contain zero, and are closed with respect tu successorship:

$$\forall x \ \forall Y \ \big((Y0 \land \forall z \ (Yz \to Ysz) \to Yx\big)$$

6 Second-Order Peano Arithmetic

When we introduced the axioms of PA, we faced some difficulty in expressing the induction principle, which says

(IP) If a property is true of zero, and it is inherited from every number to its successor, then it is true of every number.

The best way to formalize it in first-order logic was in the form of a formula scheme:

$$(\text{PA3}_1) \quad \big(\varphi(0) \land \forall x \ (\varphi(x) \to \varphi(sx))\big) \to \forall x \ \varphi(x)$$

In second-order logic we can quantify over properties, so we can express the principle in a single formula:

$$(\text{PA3}_2) \quad \forall Z \ \big(\big(Z(0) \land \forall x \ (Z(x) \to Z(sx))\big) \to \forall x \ Z(x)\big)$$

However similar these versions may seem, there's a huge difference between them. PA3_1 has as many instances as many formulas there are in the language. This is the smallest infinite cardinality; and it is called countable. Now the instances of PR3_2 are all sets of natural numbers (as extensions of properties of numbers); and it is a well-known fact of set theory that every set has more subsets than elements. Thus, since the set of natural numbers is infinite, there have to be more subsets of numbers than formulas about numbers. In other words, most sets of numbers are not parametrically definable by a first-order formula in PA. The second-order version is considerably stronger than the first-order one.

Note that because of the law of comprehension PA3_1 follows from PA3_2; so we can still use open formulas in proofs by arithmetic induction.

The result of replacing the induction scheme (PA3_1) by (PA3_2) in the system PA is called second-order Peano Arithmetic; in short, PA_2. It is a much stronger theory than PA; in fact, it is categorical. The nonstandard models of PA that we saw previously do not model PA_2.

We are not in a position to fully prove PA_2's categoricity; but we will prove that in every model of the system, every number is a standard number; there are no infinitely large numbers such that an infinite descending chain of predecessors would start from them.

Let $\varphi(x)$ be the following formula:

$$\forall Y \ \big((Y0 \wedge \forall z \ (Yz \rightarrow Ysz) \rightarrow Yx\big)$$

Let us see $\varphi(0)$ first:

$$\forall Y \ \big((Y0 \wedge \forall z \ (Yz \rightarrow Ysz) \rightarrow Y0\big)$$

This is a logical truth; and so is $\forall X \ (\varphi(x) \rightarrow \varphi(sx))$. Thus, the property of being a standard number is true of zero, and it is inherited from every number to its successor; therefore, it is true of every number. In PA_2-models there are no nonstandard numbers.

It is an interesting feature of PA_2 that the arithmetic operations are definable, as the smallest sets of triples that satisfy the axioms. We construct the definition of addition; the definition of multiplication goes the same manner. Since we didn't introduce function variables, we have to define addition as a relation, and use the axioms of the relational version of PA; but for the sake of simplicity we take 0 to be a constant, and s to be a function symbol. Let us use the following abbreviations:

$$\mathcal{A}_8(U) \ \stackrel{\text{def}}{\Leftrightarrow} \ \forall x' \ \forall y' \ \exists z' \ U(x', y', z')$$

$$\mathcal{A}_9(U) \ \stackrel{\text{def}}{\Leftrightarrow} \ \forall x' \ \forall y' \ \forall z' \ \forall z'' \ \big(\big(U(x', y', z') \wedge U(x', y', z'')\big) \rightarrow z' = z''\big)$$

$$\mathcal{A}_{10}(U) \ \stackrel{\text{def}}{\Leftrightarrow} \ \forall x' \ U(x', 0, x')$$

$$\mathcal{A}_{11}(U) \ \stackrel{\text{def}}{\Leftrightarrow} \ \forall x' \ \forall y' \ \forall y'' \ \forall z' \ \forall z'' \ \big(U(x', y', z') \rightarrow U(x', sy', sz')\big)$$

Now the definition is

$$+(x, y, z) \ \stackrel{\text{def}}{\Leftrightarrow} \ \forall U \ \big(\big(\mathcal{A}_8(U) \wedge \mathcal{A}_9(U) \wedge \mathcal{A}_{10}(U) \wedge \mathcal{A}_{11}(U)\big) \rightarrow U(x, y, z)\big)$$

7 Metalogical Properties

In the discussion of first-order Peano-arithmetic we saw a very important argument that proved the existence of nonstandard models for arithmetic. Let us repeat it here.

Let Γ be our favorite first-order theory of arithmetic; let a be an additional individual constant; and let Σ be the following set of formulas:

$0 < a$

$s0 < a$

$ss0 < a$

$\ldots$

Let us assume Γ has at least one model; and let $\mathcal{A}$ be an arbitrary one. $\mathcal{A}$ will also be a model of any finite part Σ' of Σ; and any finite part Γ' of Γ; hence it will also be a model of any finite part $\Gamma' \cup \Sigma'$ of $\Gamma \cup \Sigma$. But according to the *compactness theorem*, if any finite part of a set of formulas has a model, the whole set has a model, too. If Γ has a model, then so has $\Gamma \cup \Sigma$. But $\Gamma \cup \Sigma$ has nonstandard models only.

This argument must fail in SOL, because there PA_2 doesn't have nonstandard models. Therefore, we can conclude, the compactness theorem doesn't apply to SOL. But the compactness theorem was an obvious consequence of the completeness theorem for the first-order calculus. Indeed; if all theorems of PA_2 had a proof in a deduction system, then they would also be provable in a finite part of PA_2, too. So, we can conclude that if PA_2 has a model, there is no adequate calculus for second-order logic.

This result is of major importance. We proved it from the rather strong premise that PA_2 is a satisfiable theory. It will be nice to see that this premise is not necessary.

First we show that in any first-order language if a theory has arbitrarily large finite models, then it also has an infinite model. Let Γ be such a theory; and let us consider the following set Σ of formulas:

$$\exists x_1\ \exists x_2\ x_1 \neq x_2$$

$$\exists x_1\ \exists x_2\ \exists x_3\ (x_1 \neq x_2 \wedge x_1 \neq x_3 \wedge x_2 \neq x_3)$$

$$\ldots$$

$$\exists x_1\ \ldots \exists x_n\ (x_1 \neq x_2 \wedge x_1 \neq x_3 \wedge \ldots \wedge x_{n-1} \neq x_n)$$

$$\ldots$$

Since Γ has arbitrarily large finite models, any finite part of $\Gamma \cup \Sigma$ has a model; so by compactness $\Gamma \cup \Sigma$ itself has a model, too. But this model has to be infinite. We can conclude that finiteness of the domain is not expressible in first-order logic.

Now consider the formula

$$\neg\, \exists Y_1\ \exists Z_2\ \big(\forall x\ (Yx \rightarrow \exists! x'\ xZx') \wedge \forall x'\ \exists! x\ (Yx \wedge xZx') \wedge \exists x\ \neg\, Yx\big)$$

This formula says that the domain has no proper subset Y such that it is equinumerous with the domain; that is, there is a one-to-one correspondence Z between them. This is exactly Dedekind's version of the definition of finiteness. This formula is satisfied by arbitrarily large finite domains, but no infinite one. Which, once again, violates the compactness theorem; and therefore the compactnes theorem, too. (For the sake of simplicity we assumed the axiom of choice in the formulation; but this is not necessary.)

There are two further important consequences of the completeness theorem that fail in second-order logic; they are the downward and upward Lwenheim–Skolem theorem. The downward part claims that if a theory has an infinite model, then it also has a countably infinite model; and the upward part claims that if a theory has an infinite model, then it also has arbitrarily large infinite models. Both theorems will be violated by a second-order formula (in the minimal language) which asserts that the domain has exactly continuum many models. We skip the details of the construction, but we give the main lines of it. The formula would say that

- there is a Y such that with some element x in Y az zero and some relation S as successorship, Y models PA_2;

- there is a relation R which witnesses that the domain has exactly as many elements as many subsets Y has; that is, for every subset Y' of Y, there is an element x in the domain such that for any x', $x'Zx$ if and only if x' is in Y'.

To sum up: the categoricity of second-order theories comes at a prize; SOL lacks some of the most important metalogical properties of FOL which make the latter one a reliable and effective framework for the foundation of mathematics, and theoretical science in general.

8 SOL vs Set Theory

...

9 Ontological Commitments

...