THE JOURNAL OF PHILOSOPHY

VOLUME LXXII, NO. 16, SEPTEMBER 18, 1975

ON SECOND-ORDER LOGIC *

I SHALL discuss some of the relations between second-order logic, first-order logic, and set theory. I am interested in two quasi-terminological questions, viz., the extent to which second-order logic is (or is to be counted as) logic, and the extent to which it is set theory. It is of little significance whether second-order logic may bear the (honorific) label 'logic' or must bear 'set theory'. What matter, of course, are the reasons that can be given on either side. It seems to be commonly supposed that the arguments of Quine and others for not regarding second- (and higher-) order logic as logic are decisive, and it is against this view that I want to argue here. I shall be concerned mainly with Quine's critique of second-order logic and with some of the reasons that can be offered in support of applying neither, one, or both of the terms 'logic' and 'set theory' to second-order logic.[1]

The first of Quine's animadversions upon second-order logic that I shall discuss is to be found in the section of his *Philosophy of Logic*[2] called "Set Theory in Sheep's Clothing." Much of this section is devoted to dispelling two confusions which we can easily agree with

* I am grateful to Richard Cartwright, Oswaldo Chateaubriand, Fred Katz, and James Thomson for helpful criticism.

[1] My motive in taking up this issue is that there is a way of associating a truth of second-order logic with each truth of arithmetic; this association can plausibly be regarded as a "reduction" of arithmetic to set theory. [It is described in Chapter 18 of *Computability and Logic* by Richard Jeffrey and myself (New York: Cambridge, 1974).] I am inclined to think that the existence of this association is the heart of the best case that can be made for logicism and that unless second-order logic has *some* claim to be regarded as logic, logicism must be considered to have failed totally. I see the reasons offered in this paper on behalf of this claim as part of a partial vindication of the logicist thesis. I don't believe we yet have an assessment that is as just as it could be of the extent to which Frege, Dedekind, and Russell succeeded in showing logic to be the ground of mathematical truth.

[2] Englewood Cliffs, N. J.: Prentice-Hall, 1970; parenthetical page references to Quine are to this book.

Quine in deploring: that of supposing that '(∃*F*)' and '(*F*)' say that some (all) predicates (i.e., predicate-expressions) are thus and so, and that of supposing that quantification over attributes has relevant ontological advantages over quantification over sets. What I wish to dispute is his assertion that the use of predicate letters as quantifiable variables is to be deplored, even when the values of those variables are sets, on the ground that predicates are not *names* of their extensions. Quine writes, "Predicates have attributes as their 'intensions' or meanings (or would if there were attributes) and they have sets as their extensions; but they are names of neither. Variables eligible for quantification therefore do not belong in predicate positions. They belong in name positions" (67).

Let us grant that predicates are not names. Why must we then suppose, as the "therefore" in Quine's sentence would indicate we must, that variables eligible for quantification do not belong in predicate positions? Quine earlier (66/7) gives this argument:

> Consider first some ordinary quantifications: '(∃*x*)(*x* walks)', '(*x*)(*x* walks)', '(∃*x*)(*x* is prime)'. The open sentence after the quantifier shows '*x*' in a position where a name could stand; a name of a walker, for instance, or of a prime number. The quantifications do not mean that names walk or are prime; what are said to walk or to be prime are things that could be named *by* names in those positions. To put the predicate letter '*F*' in a quantifier, then, is to treat predicate positions suddenly as name positions, and hence to treat predicates as names of entities of some sort. The quantifier '(∃*F*)' or '(*F*)' says not that some or all predicates are thus and so, but that some or all entities of the sort named by predicates are thus and so.

If Quine had argued:

> Consider some extraordinary quantifications: '(∃*F*)(Aristotle *F*)', '(*F*)(Aristotle *F*)', '(∃*F*)(17 *F*)'. The open sentence after the quantifier shows '*F*' in a position where a predicate could stand; a predicate with an extension in which Aristotle, for instance, or 17 might be. The quantifications do not mean that Aristotle or 17 are in predicates; what Aristotle or 17 are said to be in are things that could be had *by* predicates in those positions. To put the variable '*x*' in a quantifier, then, is to treat name positions suddenly as predicate positions, and hence to treat names as predicates with extensions of some sort. The quantifier '(∃*x*)' or '(*x*)' says not that some or all names are thus and so, but that some or all extensions of the sort had by names are thus and so.

we should have wanted to say that the last two statements were false and did not follow from what preceded them. It seems to me that the same ought to be said about the argument Quine actually gives.

To put '*F*' in a quantifier may be to treat '*F*' as having a *range*, but it need not be to treat predicate positions as name positions nor to treat predicates as names of entities of any sort. Quine seems to suppose that because a variable of the more ordinary sort, an individual variable, always occurs in positions where a name but not a predicate could occur, the same must hold for every sort of variable. We may grant that the ordinary quantifications mean what Quine says they mean. But we are not thereby committed to any paraphrase containing 'name' (or any of its cognates) that purports to give the meaning of our extraordinary quantifications. Perhaps someone might suppose that variables must always *name* the objects in their range, albeit only "indefinitely" or "temporarily." However, we have no reason not to think that there might be a sort of variable, a predicate variable, that ranges over the objects in its range (these will be extensions) but does not *name* them "indefinitely" or any other way; rather, predicate variables will *have* them "indefinitely," as (constant) predicates have their extensions "definitely." Such variables would not be names of any sort, not even "indefinite" ones, but would have a range containing those objects (extensions) which could be had by predicates in predicate positions.

It may be that a suggestion is lurking that an adequate referential account of the truth conditions of sentences cannot be given unless it is supposed that all variables act as names that (indefinitely) name the objects in their range. But this is not the case. Although variables must have a range containing suitable objects, it need not be that variables of every sort indefinitely name the objects in their ranges. '$(\exists F)$' does not have to be taken as saying that some entities of the sort named by predicates are thus and so; it can be taken to say that some of the entities (extensions) had by predicates contain thus and such. So some variables eligible for quantification might well belong in predicate positions and not in name positions. And taking '*Fx*' to be true if and only if that which '*x*' names is in the extension of '*F*' in no way commits us to supposing that '*F*' names anything at all.

In the same section of *Philosophy of Logic* Quine has some advice for the logician who wants to admit sets as values of quantifiable variables and also wants distinctive variables for sets. The logician should not, Quine says, write '*Fx*' and thereupon quantify on '*F*', but should instead write '$x \in \alpha$' and then, if he wishes, quantify on 'α'. The advantage of the new notation is thought to be its greater explicitness about the set-theoretic presuppositions of second-order logic. There is an important distinction between first- and second-order logic with regard to those presuppositions, which may be part of the reason Quine insists on regarding 'F', 'G', etc. in first-order formulas as schematic letters and not quantifiable variables. In order to give a theory of truth for a first-order language which is materially adequate (in Tarski's sense) and in which such laws of truth as "The existential quantification of a true sentence is true" can be proved, it is not necessary to assume that the predicates of the language have extensions, although it does appear to be necessary to make this assumption in order to give such a theory for a second-order language.

There are reasons for not taking Quine's advice, however. One is that the notation Quine recommends abandoning represents certain aspects of logical form in a most striking way.[3] Another, and more important, reason is that the usual conventions about the use of special variables like 'α' guarantee that rewriting second-order formulas in Quine's way can result in the loss of validity or implication. For example, '$\exists F \forall x\, Fx$' is valid, but '$\exists \alpha \forall x\, x \in \alpha$' is not; and '$x = z$' is implied by '$\forall Y(Yx \rightarrow Yz)$' but not by '$\forall \alpha(x \in \alpha \rightarrow z \in \alpha)$'.

Quine disparages second-order logic in two further ways: reading him, one gets the sense of a culpable involvement with Russell's paradox and of a lack of forthrightness about its existential commitments. "This hypothesis itself viz., '$(\exists y)(x)(x \in y \equiv Fx)$' falls dangerously out of sight in the so-called higher-order predicate calculus. It becomes '$(\exists G)(x)(Gx \equiv Fx)$', and thus evidently follows from the genuinely logical triviality '$(x)(Fx \equiv Fx)$' by an elementary logical inference. Set theory's staggering existential assumptions are cunningly hidden now in the tacit shift from schematic predicate letter to quantifiable set variable" (68). Quine, of course, does not assert that higher-order predicate calculi are inconsistent. But even if they are consistent,[4] the validity of '$\exists X \forall x(Xx \leftrightarrow \sim x \in x)$', which certainly looks contradictory, would at any rate seem to demonstrate

[3] For instance, writing out the definition of the ancestral aR_*b in this notation:

$$\forall F(\forall x(aRx \rightarrow Fx) \mathbin{\&} \forall x \forall y(Fx \mathbin{\&} xRy \rightarrow Fy) \rightarrow Fb)$$

shows it to be obtained from an ordinary first-order formula by prefixing a universal quantifier, and suggests an interesting question: Is there an *existential* quantification of a first-order formula that is a satisfactory definition of the ancestral? (The answer is no.)

[4] Gentzen showed that the problem of their consistency had a very easy positive solution. See "Die Widerspruchsfreiheit der Stufenlogik," *Mathematische Zeitschrift*, XLI, 3 (1936): 357–366. An English translation, "The Consistency of the Simple Theory of Types," is contained in M. E. Szabo, ed., *The Collected Papers of Gerhard Gentzen* (Amsterdam: North-Holland, 1969).

that their existence assumptions must be regarded as "vast." A problem now arises: although '$\exists X\ \exists x\ Xx$' and '$\exists X\ \forall x\ Xx$' are also valid, '$\exists X\ \exists x\ \exists y(Xx\ \&\ Xy\ \&\ x \neq y)$' is not valid; it would thus seem that, despite its affinities with set theory and its vast commitments, second-order logic is not committed to the existence of even a two-membered set. Both of these difficulties, it seems to me, can be resolved by examining the notion of validity in second-order logic. This examination seems to show a certain surprising weakness in second-order logic.

When is a sentence valid in second-order logic? When it is true under all its interpretations. When does it follow from others? When it is true under all its interpretations under which all the others are true. What, then, is an interpretation of a second-order sentence? If we are considering "standard" second-order logic in which second-order quantifiers are regarded as ranging over *all* subsets of, or relations on, the range of the first-order quantifiers,[5] we may answer: exactly the same sort of thing an interpretation of a first-order sentence is, viz., an ordered pair of a non-empty set D and an assignment of a function to each nonlogical constant in the sentence. The domain of the function is the set of all n-tuples of members of D if the constant is of degree n, and the range is a subset of D if the constant is a function constant and a subset of $\{T, F\}$ if it is a predicate constant. [Names (sentence letters) are function (predicate) constants of degree 0; functions from the set of all 0-tuples of members of D into an arbitrary set E are of course members of E.] We need not explicitly mention separate ranges for the second-order variables that may occur in the sentence. An existentially quantified sentence $\exists\alpha\ F(\alpha)$ is then true under an interpretation I just in case $F(\beta)$ is true under some interpretation J that differs from I (if at all) only in what it assigns to the constant β, which is presumed not to occur in $\exists\alpha\ F(\alpha)$ and presumed to be of the same logical type[6] as the variable α. The other clauses in the definition of *truth in an interpretation* are exactly as you would suppose them to be. Notice that in this account no mention is made of what sort (individual, sentential, function, or predicate) of variable α is; α may be any sort of variable at all. Notice also that, if only individual variables are allowed, the account is just a paraphrase of one standard definition of *truth in an interpretation*. The definition changes neither the conditions under which a first-order sentence is true in an interpretation nor the account of what an interpretation is, but

[5] Only "standard" or "full" second-order logic is considered in this paper.

[6] Two symbols are of the same logical type if they are of the same degree and are either both predicate symbols or both function symbols.

merely extends in the obvious way the account given in (say) Mates's *Elementary Logic*[7] or Jeffrey's *Formal Logic*[8] to cover the new sorts of quantified sentences that arise in second-order logic. Quine has stressed the discontinuities between first- and second-order logic so emphatically and for so long that the obvious and striking continuities may be forgotten. In Mates's book, for example, nineteen laws of validity are stated, of which all but one (the compactness theorem) hold for second-order logic. Thus there is a standard account of the concepts of validity and consequence for for first-order sentences, and there is an obvious, straightforward, non–ad hoc way of extending that account to second-order sentences.[9]

We can now see what is shown by the validity of

$$\exists X \, \forall x (Xx \leftrightarrow \sim x \in x).$$

First of all, the sentence *is* valid: given any I, we can always find a suitable J in which '$\forall x(Bx \leftrightarrow \sim x \in x)$' is true by assigning to 'B' the set of all objects in the domain of I that do not bear to themselves the relation that I assigns to '$\in$'. Since the domain of I is a set, one of the axioms of set theory (an *Aussonderungsaxiom*) guarantees that there will always be such a subset of the domain. But without a guarantee that there is a set of all sets, we cannot conclude from the validity of '$\exists X \, \forall x(Xx \leftrightarrow \sim x \in x)$' that there is a set of all non-self-membered sets. And we have guarantees galore that there is no set of all sets. We do, of course, land in trouble if we suppose that 'x' ranges over all sets, that 'X' ranges over all sets of objects over which 'x' ranges, and that '$\in$' has its usual meaning; for then '$\exists X \, \forall x(Xx \rightarrow \sim x \in x)$' would be false. But that it would then be false does not show it to be invalid; for there is no interpretation whose domain contains all sets.

Our difficulty is thus circumvented, but at some cost. We must insist that we mean what we say when we say that a second-order sentence is valid if true under all its interpretations, and that an interpretation is an ordered pair of a *set* and an assignment of functions to constants.

There is thus a limitation on the use of second-order logic to which first-order logic is not subject. Examples such as '$\exists X \, \forall x(Xx \leftrightarrow \sim x \in x)$'

[7] 2d ed., New York: Oxford, 1972.

[8] New York: McGraw-Hill, 1967.

[9] In Part IV of *Methods of Logic*, 3d ed. (New York: Holt, Rinehart & Winston, 1972), Quine extends the notion of validity to first-order sentences with identity and discusses higher-order logic at length, but does not describe the extension of the notion of validity to second-order logic.

and '$\exists X\, \forall x\, Xx$', both valid, seem to show that it is impermissible to use the notation of second-order logic in the formalization of discourse about certain sorts of objects, such as sets or ordinals, in case there is no *set* to which all the objects of that sort belong. This restriction does not apply, as it appears, to first-order logic: ZF (Zermelo-Fraenkel set theory) is couched in the notation of first-order logic, and the quantifiers in the sentences expressing the theorems of the theory are presumed to range over all sets, even though (if ZF is right) there is no set to which all sets belong. In the case of '$\exists X\, \forall x\, Xx$', we cannot assume, for example, that the quantifier '$\forall x$' ranges over all ordinals, for then '$\exists X\, \forall x\, Xx$' would be true iff there were a set to which all ordinals belong, and there is no such set. Nor can we assume that it ranges over all the sets that there are, for it would then be true iff there were a set of all sets. Thus if we wish (as we do) to maintain that both sentences are true (because valid) and also wish to preserve the standard account of the conditions under which sentences are true, we cannot suppose that all sets belong to the range of '$\forall x$' in either, or that all ordinals belong to the range of '$\forall x$' in '$\exists X\, \forall x\, Xx$'. There is of course a step from supposing that the quantifier '$\forall x$' in '$\exists X\, \forall x\, Xx$' may not be assumed to range over all sets to supposing that all members of the range of first-order quantifiers in second-order sentences used to formalize a certain discourse must be contained in some one set (which depends upon the discourse), and there might be ways of not taking it. But all the difficulties do appear to have the same source, and seem to point to the impermissibility of second-order discourse about all sets, all ordinals, etc.

(We have been assuming all along that ZF is correct and that *sets* are the only "set-like" objects there are, the only objects to which membership is borne. If, however, as certain extensions of ZF assert, there are also certain *classes*, which are not sets, but which sets may be members of, then of course we are free to interpret '$\exists X\, \forall x\, Xx$' as saying that there is a class to which all sets belong and thus to suppose that '$\forall x$' ranges over all sets in '$\exists X\, \forall x\, Xx$'. But even if classes do exist, there is again a distinction between first- and second-order notation that is significantly like the distinction just described: we may use the former but not the latter to discuss *all members of the counterdomain* (the right field) *of* 'ϵ'. One of the lessons of Russell's paradox is that if we read 'Xx' as '(OBJECT) X bears R to (object) x', then the range of first-order quantifiers in second- but not first-order sentences may not contain all OBJECTS.)

There is a similar, but less significant, restriction on the use of the notation of first-order logic. One who uses it to formalize some discourse is committed (in the absence of special announcements to the contrary) to the non-emptiness of the ontology of the discourse and also to the presence in the ontology of references of any names that occur in the formalization. The use of names in formalization can be avoided, however, as Quine has pointed out, and various formulations of first-order logic exist in which the empty domain is permitted. But there is a striking difference between the commitment to non-emptiness of an ontology and the commitment to sethood: we believe that our own ontology is non-empty, but not that it forms a set! The contradictions appear, therefore, to teach us not that second-order logic may be inconsistent (as Quine perhaps intimates), but that it seems impossible that any "universal characteristic" should be couched in the notation of second-order logic.

What now of the existence assumptions of second- and higher-order logic, which Quine calls both "vast" and "staggering"? *Set theory* (ZF) certainly makes staggering existence *claims*, such as that there is an infinite cardinal number κ that is the κth infinite cardinal number (and hence that there is a set with that many members). Quine maintains that higher-order logic involves "outright assumption of sets the way [set theory] does."[10] Of course there are differences between set theory and higher-order logic: all set theories agree that there is a set containing at least two objects, but, as noted, '$\exists X\, \exists x\, \exists y(Xx \,\&\, Xy \,\&\, x \neq y)$' is not valid, for it is false in all one-membered interpretations. Let us try to see what the way*s* are in which second-order logic involves assuming the existence of sets.

First of all, "in second-order logic one quantifies over sets." There are certain (second-order) sentences of any given language that will be classified by second-order logic as logical truths (i.e., as valid), even though they assert, under any interpretation of the language whose domain forms a set, the existence of certain sorts of sub*sets* of the domain. (The sort depends upon the interpretation.) '$\exists X\, \forall x(Xx \leftrightarrow \sim x \in x)$' and '$\exists X\, \forall x(Xx \leftrightarrow x = x)$' are two examples. Thus, unless there exist sets of the right sorts, these sentences will be false under certain interpretations.

Now one may be of the opinion that no sentence ought to be considered as a truth of *logic* if, no matter how it is interpreted, it asserts that there are *sets* of certain sorts. Similarly, one might hold that the truth of '$\exists f\, \forall x\, Rf(x)x$' ought not to *follow* from that of

[10] *Set Theory and Its Logic*, 2d ed. (Cambridge, Mass.: Harvard, 1969), p. 258.

'$\forall x\ \exists y\ Ryx$' (even if the axiom of choice is true), or one might think that it is not *as a matter of logic* that there is a set with certain closure properties if Smith is not an ancestor of Jones (i.e., not a parent, not a grandparent, etc.).

The view that logic is "topic-neutral" is often adduced in support of this opinion: the idea is that the special sciences, such as astronomy, field theory, or set theory, have their own special subject matters, such as heavenly bodies, fields, or sets, but that logic is not about any sort of thing in particular, and, therefore, that it is no more in the province of logic to make assertions to the effect that sets of such-and-such sorts exist than to make claims about the existence of various types of planet. The subject matter of a particular science, what the science is about, is supposed to be determined by the range of the quantifiers in statements that formulate the assertions of the science; logic, however, is not supposed to have any special subject matter: there is neither any sort of thing that may not be quantified over, nor any sort that must be quantified over.

I know of no perfectly effective reply to this view. But, in the first place, one should perhaps be suspicious of the identification of subject matter and range. (Is elementary arithmetic really not *about* addition, but only *about* numbers?) And then it might be said that logic is not so "topic-neutral" as it is often made out to be: it can easily be said to be about the notions of negation, conjunction, identity, and the notions expressed by 'all' and 'some', among others (even though these notions are almost never quantified over). In the second place, unlike *planet* or *field*, the notions of *set*, *class*, *property*, *concept*, and *relation*, etc. *have* often been considered to be distinctively logical notions, probably for some such very simple reason as that anything whatsoever may belong to a set, have a property, or bear a relation. That some set- or relation-existence assertions are counted as logical truths in second- or higher-order systems does not, it seems to me, suffice to disqualify them as systems of logic, as a system would be disqualified if it classified as a truth of logic the existence of a planet with at least two satellites. Part 3 of the *Begriffsschrift*, for example, where the definition of the ancestral was first given, is as much a part of a treatise on logic as are the first two parts; the first occurrence of a second-order quantifier in the *Begriffsschrift* no more disqualifies it from that point on as a work on logic than does the earlier use of the identity sign or the negation sign. Poincaré's wisecrack, "La logique n'est plus stérile. Elle engendre la contradiction," was cruel, perhaps, but not unfair. And many of us first learned about the ancestral and other matters from a work not unreasonably entitled *Mathematical Logic*.

Another way in which second- but not first-order logic involves existential and other sorts of set-theoretic assumptions is this: via Gödelization and because of the completeness theorem, elementary arithmetic ("Z") is a suitable background theory for the development of a significant theory of validity of first-order formulas. A notion of "validity," coextensive with the usual one (truth of the universal closure in all interpretations), can be defined in the language of Z via Gödelization, and the validity of each valid formula (and no others) can then be proved in the theory, as can many general laws of validity. Moreover, the invalidity of many invalid sentences can also be demonstrated. In contrast, not only is there no hope of proving the validity of each valid second-order sentence in elementary arithmetic, the notion of second-order validity cannot even be *defined* in the language of *second-order* arithmetic. We can effectively associate with each first-order sentence a statement of arithmetic of a particularly simple form that is true if and only if the first-order sentence is valid, but no such association is even remotely possible for second-order sentences.[11] Worse, for many highly problematical statements of set theory (such as the continuum hypothesis) there exist second-order sentences that are valid if and only if those statements are true. Thus the metatheory of second-order logic is hopelessly set-theoretic, and the notion of second-order validity possesses many if not all of the epistemic debilities of the notion of set-theoretic truth.

On the other hand, although it is not hard to have some sympathy for the view that no notion of validity should be so extravagantly distant from the notion of proof, we should not forget that validity of a first-order sentence is just truth in all its interpretations. (The equation of first-order validity with provability effected by the completeness theorem would be miraculous if it weren't so familiar.) And, as we shall see below, there are notions of (first-order) logical theory which, unlike *validity*, can be adequately treated of only in a background theory that is stronger than elementary arithmetic.

While comparing set theory and second-order logic, we ought to remark in passing that the definability in set theory of the notion of second-order validity at once guarantees both the nonexistence of a reduction of the notion of set-theoretical truth to that of second-order validity and the existence of a reduction in the opposite direction: no effective—indeed no set-theoretically definable—function that assigns formulas of second-order logic to sentences of set theory

[11] There is a precise sense in which the set of valid second-order sentences is *staggeringly* undecidable: it is not definable in nth-order arithmetic, for any n. Its "Löwenheim number" is also staggeringly high.

assigns second-order logical truths to all and only the truths of set theory (otherwise set-theoretical truth would be set-theoretically definable). However, the function that assigns to each formula of second-order logic the sentence of set theory that asserts that the formula *is* a second-order logical truth reduces second-order validity to set-theoretical truth. Thus each of the notions in the series ⟨first-order validity, first-order arithmetical truth, second-order arithmetical truth, second-order validity, set-theoretical truth⟩ can always be reduced via effective functions to later ones but never to earlier ones; the notions are thus in order of increasing strength of one certain sort.

Quine writes (66) that "the logic capable of encompassing [the reduction of mathematics to logic] was logic inclusive of set theory." If second-order logic is "inclusive of set theory," it would seem to have to count as valid some nontrivial theorems of set theory, and if, among those counted as valid, there were some to the effect that certain kinds of set existed, second-order logic might seem to involve excessive ontological commitments in yet another way. And it may easily seem that second-order logic involves such commitments. For '$\exists X\, \forall x \sim Xx$' and '$\exists X\, \forall y(Xy \leftrightarrow y \subseteq x)$' are both valid and might be thought to assert that the null and power sets exist, just as all set theories say.

It seems, however, that there is a serious difficulty in supposing that *any* second-order sentence asserts, for example, that there is a set with no members; it seems that no second-order sentence asserts the same thing as any theorem of set theory, and hence that not even the smallest fragment of set theory is, in this sense, included in second-order logic.

Consider the question "What does '$\forall x\, x = x$' assert?" One may answer, "Why, that everything is identical with itself." But if one answers thus, one must realize that one's answer has a determinate sense only if the reference (range) of 'everything' is fixed. A more cautious answer might be "Why, that everything in the domain (whatever the domain may be) is identical with itself." If the natural numbers are in question '$\forall x\, \exists y\, y < x$' is false; if the rationals, true. (It seems to me that the ordinary Peano-Russell notation is less than ideal in not representing in a sufficiently vivid way the partial dependence of truth-value upon domain. In some ways it would be nicer if each quantifier were required to wear a subscript that indicated its range. It seems that the design of standard notation is influenced by the archaic view that logic is about some one fixed domain of *objects* or *individuals*, and that a logical truth is a sentence that is true no matter what relations on that domain are assigned to the predicate letters in the sentence.)

Thus the correct answer to the question, "What does '$\exists X \forall x \sim Xx$' assert?" would seem to be something like "That depends upon what the domain is supposed to be (and also upon how that domain is 'given' or 'described'). But, whatever the domain may be, '$\exists X \forall x \sim Xx$' will assert that there is a subset of the domain to which none of its members belong."

It should now appear that no valid second-order sentence can assert the same thing as any theorem of set theory. For a second-order sentence, whether valid or not, asserts something only with respect to an interpretation, whose domain may not be taken to contain all sets. But if the sentence were to assert what any particular set-theoretic statement asserts, its domain, it would seem, would have to contain all sets. '$\exists X \forall x Xx$' is valid, but does not assert that there is a universal set, which, if ZF is correct, is false; rather, it asserts that there is a subset of the domain (whichever set that may be) to which everything in the domain belongs. The quantifiers in the first-order sentences that express the assertions of ZF range over objects that do not together constitute a set. We have argued that the ranges of the variables in second-order sentences must be sets. If so, it is hard to see how any second-order sentence could express or assert what any theorem of ZF does, or that second-order logic counts as valid some significant theorems of set theory.

There is a clear sense, however, in which second-order logic can at least be said to be committed to the assertion that an empty set exists. For since the empty set is a subset of the domain of every interpretation whatsoever and is the only set to which no members of any domain belong, '$\exists X \forall x \sim Xx$' may be taken to assert the existence of the empty set independently of any interpretation, and second-order logic may thus be regarded as committed to its existence too. Moreover, higher- and higher-order logics will be committed in the same way to more and more sets.[12] In the case of second-order logic, though, the commitment is exceedingly modest; the null set is the only set to whose existence second-order logic can be said to be committed.

One sense, already noted, in which the use of second- but not first-order logic commits one to the existence of sets in this: If L_1 is the first-order fragment of an interpreted second-order language L_2 whose domain D contains no sets, then there are many logical

[12] I owe this point to Oswaldo Chateaubriand.

truths of L_1 that claim the existence of objects in D with certain properties, but there are none that claim the existence of subsets of D; however, among the logical truths of L_2 there are many such: for each predicate of L_2 with one free individual variable, there is a logical truth of L_2 that asserts the existence of a subset of D that is the extension of the predicate.

We have already seen definitions of validity and consequence for second-order sentences which bring out the obvious continuity of second- with first-order logic: validity and consequences are, as always, truth in all appropriate interpretations; the definition of an interpretation remains unchanged, as does the account of the conditions under which a first-order sentence is true in an interpretation. The account needs only to be *supplemented* with new clauses for the new sorts of sentence that arise in second-order logic. The supplementation may be given in separate clauses for each new sort of quantifier, which will be perfectly analogous to those for individual quantifiers. It may also be given in a general account of the conditions under which a sentence beginning with a quantifier is true in an interpretation, which applies uniformly to all sorts of quantifier, and of which the clauses for sentences beginning with individual quantifiers are special cases. The existence of such a definition provides *a* strong reason for reckoning second-order logic as logic. We come now to a second virtue of second-order logic, the well-known superiority of its "expressive" capacity.

If we conjoin the first two "Peano postulates," replace constants by variables, and existentially close, we obtain

$$\exists z \, \exists S (\forall x \, z \neq S(x) \;\&\; \forall x \, \forall y (S(x) = S(y) \rightarrow x = y))$$

a sentence true in just those interpretations whose domains are (Dedekind) infinite. If we do the same for the induction postulate, we obtain

$$\exists z \, \exists S \, \forall x (Xz \;\&\; \forall x (Xx \rightarrow XS(x)) \rightarrow \forall x \, Xx)$$

which is true in just those interpretations with countable domains. Thus the notions of infinity and countability can be characterized (or "expressed") by second-order sentences, though not by first-order sentences (as the compactness and Skolem-Löwenheim theorems show). Although first-order logic's expressive capacity is occasionally quite surprising, there are many interesting notions such as *well-ordering*, *progression*, *ancestral*, and *identity* that cannot be characterized in first-order logic (first-order logic without '=' in the case of *identity*!), but that can be characterized in second-. And the second-order characterizations of notions like these offer a way of regarding as inconsistent certain apparently inconsistent (infinite) sets of statements, each of whose finite subsets is consistent—a way that is not available in (compact) first-order logic. Four examples of such sets are {'Smith is an ancestor of Jones', 'Smith is not a parent of Jones', 'Smith is not a grandparent of Jones', . . .}, {'It is not the case that there are infinitely many stars', 'There are at least two stars', 'There are at least three stars', . . .}, {'R is a well-ordering', 'a_1Ra_0', 'a_2Ra_1', 'a_3Ra_2', . . .}, and, of course, {'x is a natural number', 'x is not zero', 'x is not the successor of zero', . . .}.[13]

Compare these four sets with {'Not: there are at least three stars', 'Not: there are no stars', 'Not: there is exactly one star', 'Not: there are exactly two stars'} and {'R is a linear ordering', 'a_0Ra_1', 'a_1Ra_2', 'Not: a_0Ra_2'}. There is a translation into the notation of first-order logic under which the latter two sets of statements are *formally* inconsistent. Moreover, the translation, together with an explanation of the conditions under which the translations are true in interpretations, provides an important part of the explanation of the inconsistency of the two sets. One would have hoped that the same sort of thing might be possible for the four former sets. It seems impossible, on reflection, that all the statements in any one of these four sets should be true; it also seems that the reasons for this impossibility would have to be of the same character as those which explain the inconsistency of the latter two sets, the kind of reason it has always been the business of logic to give. That the logic taught in standard courses demonstrably cannot represent the inconsistency of our four sets of sentences shows not that they are consistent after all, but that not all (logical) inconsistencies are representable by means of that logic. One may suspect that the second-order account of these inconsistencies is not the "correct" account and that perhaps some sort of infinitary logic might more accurately reflect the logical form of the sentences in question; in any event, second-order logic does not muff these cases altogether. In addition, then, to there being a "straightforward" extension of the definitions of *valid sentence* and *consequence of* from first- to second-order logic, another reason for regarding second-order logic as logic is that there are notions of a palpably logical character (*ancestral*, *identity*), which can be defined in second-order logic (but not first-) and which figure critically in inferences whose validity second-order logic (but not first-) can represent.

[13] Alfred Tarski, "On the Concept of Logical Consequence," in *Logic, Semantics, Metamathematics* (New York: Oxford, 1960), p. 410.

Let us turn now to the failure of the completeness theorem for second-order logic, which can hardly be regarded as one of second-order logic's happier features. The existence of a sound and complete axiomatic proof procedure and the effectiveness of the notion of proof guarantee that the set of valid sentences of first-order logic is effectively generable; Church's theorem shows that it is not effectively decidable. There are decidable fragments of first-order logic, e.g., monadic logic with identity, but decidability vanishes if even a single two-place letter is allowed in quantified sentences. However, in a 1919 paper called "Untersuchungen über die Axiome des Klassenkalkuls . . ."[14] Skolem showed that the class of monadic second-order sentences, in which only individual and one-place predicate variables and constants may occur, is also decidable.

Discussing the contrast between classical first-order quantification theory and an extension of it containing "branching" quantifiers, Quine writes,

> . . . there is reason, and better reason, to feel that our previous conception of quantification . . . is not capriciously narrow. On the contrary, it determines an integrated domain of logical theory with bold and significant boundaries, designate it as we may. One manifestation of these boundaries is the following. The logic of quantification in its unsupplemented form admits of complete proof procedures for validity (90).

The extension is then noted not to admit of complete proof procedures.

> A remarkable concurrence of diverse definitions of logical truth . . . suggested to us that the logic of quantification as classically bounded is a solid and significant unity. Our present reflections on branching quantification further confirm this impression. It is at the limits of the classical logic of quantifications, then, that I would continue to draw the line between logic and mathematics (91).

Completeness cannot by itself be a sufficient reason for regarding the line between first- and second-order logic as the line between logic and mathematics. We have seen, first, that monadic logic differs from full first-order logic on the score of *decidability*, every bit as significant a property as *completeness*; we have further seen that this difference persists into second-order logic; and we have discussed at length the fact that we can extend to second-order sentences the definition of *truth in an interpretation* without change

[14] Reprinted in Th. Skolem, *Selected Works in Logic* (Oslo: Universitetsforlaget, 1970), pp. 67–101, and especially pp. 93–101.

in the notation of an *interpretation*. How, then, can the *semi-*effectiveness of the set of first-order logical truths be thought to provide much of a reason for distinguishing logic from mathematics? Why *completeness* rather than *decidability* or *interpretation*? Of course there is a big difference between second- and first-order logic; there are many. There are also big differences among various fragments of first-order logic, between second- and third-order logic, and between second-order logic and set theory.

Quine does not state that the completeness theorem by itself provides sufficient reason for drawing the line, however. Another reason, or more of the reason, is given by what he calls the "remarkable concurrence of diverse definitions of logical truth." One of these diverse definitions is the usual one: a sentence (or "schema," in Quine's terminology) is a logical truth if it is satisfied by every model, i.e., if it is true under all its interpretations. The other is that a sentence of a reasonably rich language is a logical truth if truths alone come of it by substitution of (open) sentences for its simple component sentences. The languages in question are interpreted languages (otherwise the notion of truth of a sentence of a language, used in the definition, would be incomprehensible), and their grammar has been "standardized," i.e., put into the notation of the first-order predicate calculus, *without* function signs or identity. As usual, "reasonably rich" has to do with arithmetic. For Quine's purposes, a language may be taken to be reasonably rich if its ontology contains all natural numbers (or an isomorphic copy) and its ideology contains a one-place predicate letter true of the natural numbers (or their copies) and two three-place predicate letters representing the sum and product operations.

By appealing to a generalization of Löwenheim's theorem that is due to Hilbert and Bernays—any satisfiable schema is satisfied by a model whose domain is the set of natural numbers and whose predicates are assigned relations on natural numbers *that can be defined in arithmetic*—Quine proves a result he calls remarkable: a schema is provable (in some standard system) if and only if it is valid (true in all its interpretations), if and only if every substitution instance of it in any given reasonably rich object language is true. Dually, a schema is irrefutable if and only if it is satisfiable (true in at least one interpretation), if and only if some substitution instance of it in the object language is true. (The equivalence of validity and provability, and of satisfiability and irrefutability, is guaranteed by the completeness theorem.)

For the purposes of this theorem, Quine cannot count the identity sign as a logical symbol: '$\exists x \, \exists y \sim x = y$' is a schema and also a sentence whose only substitution instance is itself (if '=' counts as a logical symbol), which is true (since there exist at least two objects in the domain of the object language), but which is not a logical truth according to the usual definition, for it is false in all one-membered interpretations.

A second minor point about the definition is that it just does not work if the object language is not reasonably rich.[16] But the language of arithmetic, interpreted in the usual way, is certainly reasonably rich, or becomes so when '+' and '·' are supplanted by three-place predicate letters.

The theorem may be remarkable, but it is not, I think, remarkably remarkable. A distinction can be drawn between two kinds of completeness theorem that can be proved about systems of logic: between weak and strong completeness theorems. A weak completeness theorem shows that a sentence is provable whenever it is valid; a strong theorem, that a sentence is provable *from a set* of sentences whenever it is a logical consequence of the set. Most of the usual proofs of the weak completeness of systems of first-order logic can be expanded quite easily to proofs of the strong completeness of those systems. The strong completeness of first-order logic can be expressed: a set of sentences is satisfiable if it lacks a refutation. (A refutation of a set of sentences is a proof of the negation of a conjunction of members of the set.)

It seems to me that the concurrence of the two accounts of the concept of logical truth cannot be called remarkably remarkable if their extensions to the relation of logical consequence do not concur. If there is a reasonably rich language and a set of sentences in that language which is satisfiable according to the usual account but which cannot be turned into a set of truths by (simultaneous, uniform) substitution of open sentences of the language, then the interest of the alternative definition of logical truth is somewhat diminished, for it is a definition that cannot be extended to kindred logical relations in the correct manner. And, as it happens, there is a satisfiable set of sentences of a reasonably rich language with this property. Proof is given in the appendix.

The compactness theorem might be thought to provide a way out of the difficulty. Since a set is satisfiable if and only if all its finite subsets are satisfiable, we might propose to define satisfiability by saying that a set is satisfiable just in case every conjunction of its members has a true substitution instance. So there turn out to be

[16] See Peter G. Hinman, Jaegwon Kim, and Stephen P. Stich, "Logical Truth Revisited," this JOURNAL, LXV, 17 (Sept. 5, 1968): 495–500.

three accounts of satisfiability of sets of sentences, the account just mentioned, truth in some one model, and irrefutability.

But this concurrence is not in the least remarkable. The strong completeness theorem is remarkable; and the Löwenheim-Hilbert-Bernays theorem is remarkable. The concurrence of the two definitions of validity of single sentences—truth in all interpretations and truth of all instances—is remarkable too, *because both definitions have some antecedent plausibility as correct explications of a pre-theoretical notion of logical validity* ("truth regardless of what the nonlogical words mean"). The definition of satisfiability of a set as "truth of some instance of each conjunction of schemata in the set" has no such plausibility as an account of satisfiability. It even sounds wrong.

One ought then to be wary of the claims that the concurrence of diverse definitions of logical truth is remarkable and that this concurrence suggests that classical quantificational logic is a "solid and significant unity." One of the definitions is a definition of logical truth only in virtue of a remarkable theorem about first-order logic; another cannot be generalized properly. Does classical quantificational logic then fail to be a significant and solid unity? Certainly not.

GEORGE S. BOOLOS

Massachusetts Institute of Technology

APPENDIX

We consider two first-order languages (without '='), L and M, whose predicate letters are F, Z, S, P, T, and G. The variables of both languages range over the natural numbers, and both specify that F is true of all natural numbers, that Z is true of zero alone, and that S, P, and T are predicate letters for successor, sum, and product, respectively. L specifies that G is true of all natural numbers. L is a reasonably rich language. Let A be the set of Gödel numbers of truths of L. A is not definable in L. Finally, M specifies that G is true of all and only the members of A.

Let B be the set of truths of M. B is satisfiable. But B cannot be turned into a set of truths of L by substitution of open sentences of L for the predicate letters F, Z, S, P, T, and G. For, if it could, A would be recursive in the extensions in L of the open sentences substituted for Z, S, and G, and hence A would be definable in L; for the extensions would certainly be definable in L, and *definable in L* is closed under *recursive in.*

Let '$E(\theta)$' abbreviate 'the extension in L of the open sentence substituted for θ'. The reason that A would be recursive in $E(Z)$, $E(S)$, $E(G)$ is that, for each natural number n,

$\exists x_0x_1\cdots x_{n-1}x_n(Zx_0 \mathbin{\&} Sx_0x_1 \mathbin{\&} \cdots \mathbin{\&} Sx_{n-1}x_n)$ is in B;

if $n \in A$, then

$\forall x_0x_1\cdots x_{n-1}x_n(Zx_0 \mathbin{\&} Sx_0x_1 \mathbin{\&} \cdots \mathbin{\&} Sx_{n-1}x_n \rightarrow Gx_n)$ is in B; and if $\sim n \in A$, then

$\forall x_0 x_1 \cdots x_{n-1} x_n (Zx_0 \,\&\, Sx_0x_1 \,\&\, \cdots \,\&\, Sx_{n-1}x_n \rightarrow \sim Gx_n)$ is in B.

Then, to determine whether $n \in A$, we may use "oracles" for $E(Z)$ and $E(S)$ to find an $(n+1)$-tuple $a_0, a_1, \cdots, a_{n-1}, a_n$ of natural numbers such that a_0 is in $E(Z)$ and the n pairs $a_0, a_1, \cdots$, and a_{n-1}, a_n are in $E(S)$, and then use an oracle for $E(G)$ to determine whether a_n is in $E(G)$. a_n is in $E(G)$ iff $n \in A$. The procedure is recursive in $E(Z)$, $E(S)$, $E(G)$.

We have thus shown that B is a satisfiable set of sentences of the reasonably rich language L which cannot be turned into a set of truths by (simultaneous, uniform) substitution of open sentences of L for the predicate letters of L which occur in the sentences in B.

NEW BOOKS IN PAPERBACK

ADAM, MICHEL: *Essai sur la bêtise.* Paris: Presses Universitaires de France, 1975. 195 p.

AHLUWALIA, JASBIR SINGH: *Marxism and Contemporary Reality.* New York: Asia Publishing House, 1973. 59 p. $5.25.

AIMONETTO, ITALO: *Le Antinomie logiche e matematiche.* Turin, Italy: Filosofia, 1975. 79 p. L 2,200.

AMBACHER, MICHEL: *Les Philosophies de la nature.* Paris: Presses Universitaires de France, 1974. 126 p.

ANSART-DOURLEN, MICHELE: *Dénaturation et violence dans la pensée de J. J. Rousseau.* Paris: Klincksieck, 1975. 302 p. F 75.

ARNHEIM, RUDOLF: *Entropy and Art: An Essay on Disorder and Order.* Berkeley: University of California Press, 1974. 64 p. (Cal 275.) $1.65.

BARION, JAKOB: *Was ist Ideologie?* Bonn: Bouvier, 1974. 164 p. DM 19,80.

BERLINGER, RUDOLPH: *Philosophie als Weltwissenschaft.* (Elementa. Band 2.) Amsterdam: Rodopi, 1975. 240 p. $20.00.

BERTMAN, MARTIN A.: *Research Guide in Philosophy.* Morristown, N.J.: General Learning Press, 1974. viii, 252 p.

BLANCHÉ, ROBERT: *L'Induction scientifique et les lois naturelles.* Paris: Presses Universitaires de France, 1975. 170 p.

BLANCHÉ, ROBERT: *Le Raisonnement.* Paris: Presses Universitaires de France, 1973. 263 p.

BLUMENTHAL, ALBERT: *Moral Responsibility: Mankind's Greatest Need. Principles and Practical Applications of Scientific Utilitarian Ethics.* Santa Ana, Calif.: Rayline Press, 1975. 500 p. $7.00.

BOOTH, WAYNE C.: *Modern Dogma and the Rhetoric of Assent.* Chicago: University Press, 1974. xvii, 235 p. $3.95.

BOSLEY, RICHARD: *Aspects of Aristotle's Logic.* New York: Humanities, 1975. 137 p. $10.00.

BRENNAN, JOSEPH GERARD: *Ethics and Morals.* New York: Harper & Row, 1973. ix, 398 p. $6.95.