THE BULLETIN OF SYMBOLIC LOGIC
Volume 4, Number 3, Sept. 1998

TRUTH DEFINITIONS, SKOLEM FUNCTIONS AND AXIOMATIC SET THEORY

JAAKKO HINTIKKA

§1. The mission of axiomatic set theory. What is set theory needed for in the foundations of mathematics? Why cannot we transact whatever foundational business we have to transact in terms of our ordinary logic without resorting to set theory? There are many possible answers, but most of them are likely to be variations of the same theme. The core area of ordinary logic is by a fairly common consent the received first-order logic. Why cannot it take care of itself? What is it that it cannot do? A large part of every answer is probably that first-order logic cannot handle its own model theory and other metatheory. For instance, a first-order language does not allow the codification of the most important semantical concept, viz. the notion of truth, for that language in that language itself, as shown already in Tarski (1935). In view of such negative results it is generally thought that one of the most important missions of set theory is to provide the wherewithal for a model theory of logic. For instance Gregory H. Moore (1994, p. 635) asserts in his encyclopedia article "Logic and set theory" that

> Set theory influenced logic, both through its semantics, by expanding the possible models of various theories and by the formal definition of a model; and through its syntax, by allowing for logical languages in which formulas can be infinite in length or in which the number of symbols is uncountable.

The obvious rejoinder here is that one can perfectly well describe models of any sort and define what it means for a structure to be a model for a given sentence without resorting to set theory. The most obvious candidates for this role are higher-order logics. Likewise, infinite formulas can be described by means of the resources of higher-order and sometimes even first-order logic. Admittedly, first-order logic alone cannot do all the jobs that set theory was reputedly introduced to do. For instance, one cannot define truth for a given first-order language in that particular first-order language, but one might be able to define it in another, richer, language. Indeed, Tarski

Received July 29, 1997; revised April 17, 1998.

1079-8986/98/0403-0003/$4.50

himself studied the conditions on which such a definition is possible. (See Tarski 1935.) For instance, as Tarski in effect showed, one can define truth for a first-order language in the corresponding second-order language.

However, before we indulge in such "semantic ascent," we should be clear about the reasons for it. In this paper, it will turn out logicians have universally missed the true, exceedingly simple feature of ordinary first-order logic that makes it incapable of accommodating its own truth predicate. (See Section 4 below.) This defect will also be shown to be easy to overcome without transcending the first-order level. This eliminates once and for all the need of set theory for the purposes of a metatheory of logic.

Behind the idea that we need set theory for a metatheory of logic, there thus often lurks the assumption that only first-order logic is really logic and that higher-order logic is, as Quine once put it, "set theory in sheep's clothing." Ironically, the selfsame inability of ordinary first-order logic to serve as its own metatheory has been used by Hilary Putnam (1971) as an argument against the identification of first-order logic as *the* logic. Thus there is obviously a great deal of uncertainty and confusion about the relationship of higher-order logics to set theory and indeed about the status of these logics. It has for instance, been argued in effect by Michael Friedman (1988) that the inadequacy of ordinary first-order logic as its own model theory is the fundamental reason why Carnaps's project of a general "logical syntax of language" was doomed to fail. What will be established in this paper shows that both Putnam's and Friedman's lines of argument are inconclusive.

Be its motivation what it is, set theory is often considered as some sort of instant model theory—or perhaps we should say, non-Californians' model theory. For instance, when the sharpest philosophers of science realized that a study of "the logical syntax of the language of science" was not enough, they resorted to set theory for their conceptualizations. Ironically some misguided philosophers of science have continued to seek salvation in set theory even long after the development of logical semantics and systematic model theory.

In a different corner of the philosophical world, Wittgenstein's fanatical hatred of set theory can only be understood as a corollary to his deep-seated and total rejection of all model-theoretical and other metatheoretical conceptualizations.

Such ideas of set theory as common folks' model theory cannot be shrugged off as popular misconceptions. It is hard to see what foundational interest set theory would have if it could not serve as a universal framework of all model theory. It even seems likely (as Moore apparently suggests) that part of the early motivation of set theory was to see it as a universal source of models for all and sundry theories, including those that cannot be interpreted in already known theories. Such a conception of set theoretical universe as "the model of all models" seems to lurk in wings of much of the early twentieth-century foundational discussion. It is even present in Tarski's classical monograph (Tarski 1935). Contrary to what is often said and thought, Tarski does not there show how to define truth in a structure (model). He shows how to define it in *the* model. He is there envisaging only one omnicomprehensive structure of structures of which for instance the models of the theories Hilbert's school were studying form a subset (Tarski 1956, p. 199).

In order to avoid misunderstandings it is important to realize that there is a set theory and there is a set theory. Conceived of as a study of different kinds of infinite cardinals and ordinals, set theory is of course unobjectionable, but it is then only one mathematical theory among others, without a claim to be a universal framework of all model theory.

Now the current incarnation of the idea of general set theory is axiomatic set theory. It exists in different varieties, but the differences, say, between the Zermelo-Fraenkel set theory and the von Neumann-Bernays set theory are not relevant to our purposes in this paper. What is crucial is that they are first-order theories in the sense of employing ordinary first-order logic as their sole logical component. We also have to make the minimal assumption that the set theory in question is rich enough for us to do elementary arithmetic in it.

But if axiomatic set theory is supposed to be an implementation of the idea of set theory as poor man's model theory, tables can be turned on it. It will be shown that truth can after all be defined for a suitable first-order language in that language itself, but that truth in a model of axiomatic set theory cannot be defined in that set theory. Moreover, this failure is not just a corollary to the deductive incompleteness of axiomatic set theory, which has necessitated a search for new, stronger axioms. No, it will be shown that any attempt to define set-theoretical truth in axiomatic set theory yields wrong results. Our argument does not turn on any requirement, either, that the models of axiomatic set theory be standard in something like Henkin's (1960) sense.

Since truth is the central concept of all semantics, it will thus be argued that, while at least some of the central parts of the model theory (semantics) of first-order logic can be expressed in a first-order language, the semantics of axiomatic set theory cannot in general be dealt with in such a set theory. It is hard not to consider this as a failure of its original mission.

§**2. Truth predicates for arithmetical languages.** How can all this be shown? The positive part is relatively easy. Since essentially the same story has been told elsewhere, perhaps we can afford to be brief. We will consider first a first-order arithmetical language L—more generally, a first-order language that includes elementary number theory—which has as its twin the corresponding second-order language $L^{(2)}$. In such a first-order language we can carry out a Gödel numbering for all formulas both for that first-order language L and for its second-order extension $L^{(2)}$. Actually, we can restrict our attention to the $\Sigma_1^1\mathcal{M}$ fragment $L^{(*)}$ of $L^{(2)}$. It is assumed that the logical constants of L are $=, \sim, \&, \vee, (\exists x), (\forall y)$.

How can we set up a truth-definition for L in $L^{(*)}$? We will use an idea that perhaps can be considered as a mirror image of Tarski's (1956, Section 3) recursive method of truth-definition. We will consider the characteristic properties that any truth predicate say X will have to satisfy. (This X will of course be a complex numerical predicate of natural numbers, including Gödel numbers of formulas.) If we succeed in expressing enough of those characteristics to make sure that it behaves in the right way, we can then attribute truth to the Gödel number $g(S)$ of a sentence S by simply saying that there exists such a predicate and that $g(S)$ has it. If $\mathrm{Tr}[X]$ is a summary of the characteristics in question, then the truth predicate will be

(2.1) $(\exists X)(\mathrm{Tr}[X] \ \& \ X(y))$

where y ranges over natural numbers, including Gödel numbers of sentences. If $\mathrm{Tr}[X]$ is a $\Sigma_1^1\mathcal{M}$ formula, then so will obviously be (2.1).

What the different clauses of $\mathrm{Tr}[X]$ will say can be expressed by a conjunction of conditionals and biconditionals. In the following table, the left hand side describes the antecedent of each conditional while the right hand side describes the corresponding consequent.

(2.2)	$(S_1 \ \& \ S_2)$ is true	S_1 and S_2 are true

More literally, this can be expressed as follows:

(2.3)	X applies to the Gödel number of $(S_1 \ \& \ S_2)$	X applies to the Gödel numbers of S_1 and of S_2

Denoting "Gödel number of" by g, we can write:

(2.4)	$X(g(S_1 \ \& \ S_2))$	$(X(g(S_1)) \ \& \ X(g(S_2)))$

It is well known (and fairly obvious) that $g(S_1)$ and $g(S_2)$ are recursive functions of $g((S_1 \ \& \ S_2))$ and hence representable in the language of elementary arithmetic. (More generally, for questions of representability of different relations in elementary arithmetic, see any standard exposition of Gödel numbering, and Gödel's incompleteness theorem, for instance Mendelson 1987, pp. 149–168.) This means that the general conditional from

(2.5) $X(g(S_1 \ \& \ S_2))$

to

(2.6) $(X(g(S_1) \ \& \ X(g(S_2)))$

is expressible in the language in question. This conditional is one of the conjuncts in $\mathrm{Tr}[X]$.

Likewise, we can write:

(2.7) $X(g(S_1 \vee (S_2))$ $\qquad$ $X(g(S_1)) \vee (X(g(S_2))$

In arithmetical languages, one can define a function r which takes us from a number n to the Gödel number of the numeral $\mathbf{n} = r(n)$ representing it. Hence we can formulate a conditional from:

(2.8) $X(g((\exists x)S[x]))$ $\qquad$ There is a number n such that $X(g(S[\mathbf{n}]))$

This will be another conjunct in $\mathrm{Tr}[X]$. Here we are again relying on the fact that we can get from $g((\exists x)S[x])$ and n to $g(S[r(n)])$ by means of a recursive function representable in the given arithmetical language.

Likewise we can formulate a conditional from

(2.9) $X(g((\forall x)S[x]))$ $\qquad$ to $\qquad$ For every number n, $X(g(S[\mathbf{n}]))$

Of course, what (2.8)–(2.9) are intended to capture are conditionals involving the following antecedents and consequents:

(2.10) $(\exists x)S[x]$ is true $\qquad$ For some number n, $S[\mathbf{n}]$ is true
(2.11) $(\forall x)S[x]$ is true $\qquad$ For every number n, $S[\mathbf{n}]$ is true

As was indicated, all these conditionals can be expressed in L.

The easiest way of dealing with negated expressions is to add clauses that serve to push negation signs deeper into formulas. The following are examples of such clauses:

(2.12) $X(g(\sim(S_1 \,\&\, S_2))) \supset (X(g(\sim S_1)) \vee (X(g(\sim S_2))$
(2.13) $X(g(\sim(\forall x)S[x])) \supset (X(g((\exists x) \sim (S[x]))$
(2.14) $X(g(\sim\sim S)) \supset (X(g(S))$

Once again, (2.12)–(2.14) can be seen to be expressible in our arithmetical language.

Last but not least we need a clause for atomic sentences. For instance, for atomic sentences containing a two-place primitive predicate R we can use a clause that says the following:

(2.15) $X(g(R(\mathbf{a}, \mathbf{b}))) \leftrightarrow R(a, b)$

Once again, this can be expressed in the kind of language we have considered.

The formula $\mathrm{Tr}[X]$ contains such a conjunct (clause) for each primitive predicate and function of the language in question as well as for each kind of negated atomic sentence. Notice that these are the only biconditionals among the conjuncts in $T(X)$.

For instance, a substitution-instance of (2.15) might be

$$X(g(\text{“}(1+2=3)\text{”})) \leftrightarrow ((1+2=3).$$

The identity predicate can be handled in the same way as primitive non-logical predicates, and so can negated atomic sentences.

This completes the definition of our truth predicate. It defines truth for the Gödel numbers of all the sentences of L in $L^{(*)}$, for clearly our truth predicate is a $\Sigma_1^1\mathcal{M}$ formula. According to a well-known result (see e.g., Walkoe 1970), it can therefore be expressed in the corresponding independence-friendly (IF) *first-order* language. (These languages are presented in Hintikka 1996, chapters 3–4 and in Hintikka and Sandu 1997.) What is characteristic of IF languages is that they permit an existential quantifier $(\exists y)$ to be independent of a universal quantifier $(\forall x)$ within whose syntactical scope it occurs. This is indicated by writing it $(\exists y/\forall x)$. Similarly, any quantifier and propositional connectives can be liberated from the scope of other quantifiers and/or connectives. Independence in the intended sense is in a game-theoretical semantics manifested simply as informational independence in the sense of general game theory: a move (in this case, a choice of the value of a quantified variable) is made in ignorance of a certain earlier move.

Needless to say, the truth predicate defined above cannot be expressed by the sole means of ordinary first-order logic. This is probably the reason why Tarski did not pursue the line of thought followed here. He realized that a truth definition could only be expressed in another, stronger language. Since he knew only the ordinary ones, he did not formulate his truth definitions on the first-order level.

Similar truth-predicates can be defined in other languages. In order to keep the technicalities at a minimum, we will not present the full details here, however.

§3. Truth predicates, Skolem functions and the excluded middle. The most important fact about the truth predicate just defined is the following:

> *For any sentence S of L, the truth predicate applies to $g(S)$ if and only if all the Skolem functions of S exist.*

An explanation is needed here. By the Skolem functions of a first-order formula, we mean the choice functions for its several existential quantifiers. Thus the Skolem function form of any formula S_0 in the negation normal form is obtained by replacing each existentially quantified subformula $(\exists x)S_1[x]$ of S_0 , that is, of

(3.1) $\quad S_0[(\exists x)S_1[x]]$

by $S_1[f(y_1, y_2, \dots)]$, so that (3.1) becomes

(3.2) $\quad S_0[S_1[f(y_1, y_2, \dots)]]$

where $(\forall y_1)$, $(\forall y_2)$, $\dots$ are all the universal quantifiers within the scope of which $(\exists x)S_1[x]$ occurs in S_0 and f is a new function symbol, different for different existentially quantified formulas $(\exists x)S_1[x]$.

For the purposes of this paper, we will extend the notion of Skolem function to cover also choice functions associated with disjunctions. Thus, if a disjunction $(S_1 \vee S_2)$ occurs as a subformula of S_0 that is, of

(3.3) $S_0[(S_1 \vee S_2)]$,

then in the (extended) Skolem function form it is replaced by

(3.4) $((S_1 \,\&\, (h(y_1, y_2, \dots) = 0)) \vee (S_2 \,\&\, (h(y_1, y_2, \dots) \neq 0)))$

where $(\forall y_1), (\forall y_2), \dots$ are all the universal quantifiers within the scope of which $(S_1 \vee S_2)$ occurs in S_0 and h is a different function symbol for each disjunction. (The functions h are assumed to be different from the functions f.)

To say that the Skolem functions for S_0 exist is to assert the result of generalizing existentially with respect to all the functions f, h.

The result formulated above in terms of Skolem functions will be called the *Skolemization Lemma*. It can be proved most naturally by considering the following three statements about a given sentence S of L:

(i) The truth predicate (2.11) applies to $g(S)$.
(ii) S is true in the game-theoretical sense.
(iii) The Skolem functions of S exist.

Here it is easy to prove the equivalences (i) $\to$ (ii), (ii) $\to$ (iii), and (iii) $\to$ (i). In order to prove them, we nevertheless have to explain what the game-theoretical sense of truth is. This notion of truth is defined by reference to certain two-person zero-sum games $G(S)$ associated with first-order sentences S, given a model M with the domain of individuals $\mathrm{do}(M)$ on which the nonlogical primitives of our language have been interpreted. There are two players, the verifier and the falsifier. The game rules are as follows:

(3.5) $G(S_1 \,\&\, S_2)$ begins with a choice of S_i ($i = 1$ or 2) by the falsifier. The game is then continued as in $G(S_i)$.

(3.6) $G(S_1 \vee S_2)$ begins with a choice of S_i ($i = 1$ or 2) by the verifier. The game is then continued as in $G(S_i)$.

(3.7) $G((\exists x)S[x])$ begins with a choice of $b \in \mathrm{do}(M)$ by the verifier. The game is then continued as in $G(S[\mathbf{b}])$.

It has to be assumed here that each individual $b \in \mathrm{do}(M)$ has a name $\mathbf{b} = r(b)$ whose Gödel number is a function of b expressible in the language.

(3.8) $G((\forall x)S[x])$ likewise, except that the choice is made by the falsifier.

(3.9) Negation can be dealt with by pushing the negation signs as deep into the formulas as they can go.

After a finite number of moves a game will reach a negated or unnegated atomic sentence. Since the primitives of our language have been interpreted on $\mathrm{do}(M)$, this output sentence is true or false. If true the verifier has won. If false, the falsifier has won.

(3.10) The sentence S is defined to be true iff there exists a winning strategy for the verifier, false if and only if there is a winning strategy for the falsifier.

Now the implications needed for the proof of the Skolemization Lemma are virtually obvious. For the Skolem functions of S codify the crucial parts of the verifier's strategy and that strategy clearly wins if and only if S is true.

The Skolemization Lemma shows that there is a close connection between first-order truth-definitions and Skolem functions. This connection will be the focus of this paper.

Another important general result is the following:

If there are no quantifiers in the closed sentence S, then $\mathrm{Tr}[g(S)] \leftrightarrow S$.

I will call this result the T-theorem because of its similarity with Tarski's T-schema. This formulation is not adequate, however. Usually S contains numerical constants $n_1, n_2, \ldots$. They are represented in the formalism by the numerals $r(n_1), r(n_2), \ldots$. Hence the equivalence in question should be expressed by the following:

$$\mathrm{Tr}[g(S[r(n_1), r(n_2), \ldots])] \leftrightarrow S[n_1, n_2, \ldots]$$

This can be extended to all constant numerical terms $t_1, t_2, \ldots$ over and above numerical constants $n_1, n_2, \ldots$.

The proof of this theorem is almost immediate. From the equivalence of my truth predicate with the usual truth conditions it follows that the clauses for $\&, \vee, \sim$ can be strengthened to biconditionals. After this strengthening, the truth of the T-theorem can be seen from the strengthened clauses for $\&, \vee, \sim$ plus the clauses for atomic sentences and identities.

The truth predicate defined above was first formulated as a $\Sigma_1^1\mathcal{M}$ formula. As such, it can be translated (as was indicated) into corresponding independence-friendly (IF) first-order language. (Cf. Hintikka 1996, chapter 3.) This possibility of first-order truth predicate might at first sight seem puzzling. As was emphasized already by Tarski (1935), attempted truth-definitions in terms of ordinary first-order logic normally give rise to liar-type inconsistencies. In the transition from ordinary first-order logic to IF one, the only novelty is to allow for informational independence in semantical games. This apparently increases rather than reduces what can be proved by purely logical means. How, then, can liar-type counter-examples be avoided in IF first-order languages?

The first step toward an answer is to point out that the *tertium non datur* fails in IF first-order logic. Hence liar-type paradoxes are avoided because the critical self-referential sentences constructible along Gödelian lines turn out to be neither true nor false.

Such ways of avoiding liar-type paradoxes are a dime a dozen, however, and they have been in the past criticized for arbitrariness. What sets the treatment presented here apart, is not what it assumes (arbitrarily or nonarbitrarily), but what it does not assume. From the point of view of game-theoretical semantics, the failure of the law of excluded middle happens automatically, without any new assumptions whatsoever. (Except, of course, for allowing informational independence in the sense of general theory of games.) What is more, in this approach the failure of *tertium non datur* is precisely what is to be expected. For in it truth is defined as the existence of a winning strategy for the original verifier and falsity as the existence of a winning strategy for the original falsifier in the same semantical game. Hence the principle of *tertium non datur* becomes the thesis that all semantical games are determinate. But from the theory of games it is known (if the point is not obvious in itself) that many two-person games are not determinate: neither player has a strategy that wins against all the strategies of the other. That some semantical games are not determinate should therefore come as no surprise. This indeterminacy can be considered as the underlying reason for the failure of *tertium non datur* that in turn defeats liar-type counter-examples to truth-definitions. The fact that the law of excluded middle holds in ordinary first-order logic must be considered as a fortunate but unrepresentative accident.

§**4. Direct truth definitions.** The Skolemization Lemma deserves a few comments. It is not difficult to see how the different conjunctions in the truth predicate (1) help to make sure that there exists a winning strategy for the initial verifier in the game played with the sentence whose Gödel number is y. Each conditional makes sure that a putative winning strategy can be continued by a step in one of the different types of situations which the verifier might encounter in the semantical game. It is hence a truth definition that is geared to the extensive form of semantical games.

This prompts the question: Can we formulate a truth predicate which is similarly geared to the von Neumann normal form of semantical games? At first, the task seems to be easy. According to the Skolemization Lemma, all we have to do is to say that S is true if and only if its Skolem functions exist.

In applying this idea we can even assume that a number of preliminary simplifications have been carried out. For these simplifications are obviously effective ones. By Church's thesis, the relationship of the Gödel numbers of the given sentence and that of its simplified form is recursive. (For this kind of argument, see Rogers 1967, pp. 20–21.) Hence it can be represented in the language we are assumed to be working in.

First, we will extend our discussion to IF first-order languages. In other words, L is now assumed to be an independence-friendly first-order language. In such a language, the sentence whose truth or non-truth we are examining, i.e., to whose Gödel number the truth predicate is supposed to apply or not to apply, can be assumed to be of the following form:

$$(4.1)\quad (\forall x_1)(\forall x_2)\ldots(\forall x_m)(\exists y_1/\forall x^*_{11},\forall x^*_{12},\ldots)\\ (\exists y_2/\forall x^*_{21},\forall x^*_{22},\ldots)\ldots(\exists y_n/\forall x^*_{n1},\forall x^*_{n2},\ldots)\\ S[x_1,x_2,\ldots,x_m,y_1,y_2,\ldots,y_n]$$

Here each variable set $\{x^*_{j1},x^*_{j2},\ldots\}$ is a subset of $\{x_1,x_2,\ldots,x_m\}$. We will denote by

$$(4.2)\quad \langle x_{j1},x_{j2},\ldots\rangle$$

the result of omitting from $\langle x_1,x_2,\ldots,x_m\rangle$ all the variables $x^*_{j1},x^*_{j2},\ldots$. It will be assumed that in (4.1) every disjunction of the form

$$(S_{k1}(\vee/\forall x_{j1},\forall x_{j2},\ldots)S_{k2})$$

has been replaced by

$$((S_{k1}\ \&\ (y_j=0)\vee(S_{k2}\ \&\ (y_j\neq 0))))$$

where y_j is bound to one of the initial existential quantifiers, viz., to

$$(\exists y_j/\forall x_{j1},\forall x_{j2},\ldots)$$

This eliminates all the slashes from the unquantified part of (4.1).

In view of the Skolemization Lemma, it might seem tempting to think that we can solve all our problems simply by saying that the truth condition of (4.1) is, in its $\Sigma^1_1\mathcal{M}$ form.

$$(4.3)\quad (\exists f_1)(\exists f_2)\ldots(\exists f_n)(\forall x_1)(\forall x_2)\ldots(\forall x_m)\\ S[x_1,x_2,\ldots,x_m,f_1(x_{11},x_{12},\ldots),f_2(x_{21},x_{22},\ldots),\\ \ldots,f_n(x_{n1},x_{n2},\ldots)]$$

Of course, in a sense (4.3) is a true truth condition of (4.1). But the formulation (4.3) is not addressed to the problem of formulating a truth predicate of the right sort. Such a truth predicate is supposed to be a predicate of the Gödel numbers of sentences. Hence the real question here is not to formulate truth conditions for a given sentence like (4.1) when it is written out and given in the form of its quotation. What we have to do to formulate a predicate that applies to the sentence in question even when it is not spelled out, in other words, to formulate a predicate that could apply to the Gödel number $g((4.1))$ of (4.1). Philosophical logicians have perhaps followed Tarski's exposition too closely and been led to think that sentences can be thought of as being given by means of quotations. This is not the case when we are dealing with a syntax arithmetized by means of a Gödel numbering. Then what is given to us is not a quoted expression like (4.1), but its Gödel number. (Hence when one's syntax is arithmetized explicitly it does not for instance make sense to speak of a disquotational conception of truth.) The real task in formulating the truth predicate is to show how the truth condition of a sentence like (4.1) depends on its Gödel number *and nothing else* (besides the logic we are dealing with). Only when this is done explicitly can we hope to see from that representation the ins and outs of truth-predicates for IF first-order languages.

This shows what the original sin of the disquotational theorists of truth is. They assume that we can unproblematically speak of a given sentence as being represented by its quotation. In reality, the intriguing process is how to reach the sentence in question when only its Gödel number is given, or alternatively how to reach a natural truth condition for it.

For the purpose of showing how the truth condition of (4.1)—in effect (4.3)—depends on the Gödel number g of (4.1), it can first be noted that certain relations and functions can be represented in the underlying first-order arithmetical language. They will include the following:

$\mathrm{Seq}(\alpha, m)$	$=$	α is a sequence of m numbers

It is here assumed that finite sequences of numbers can be coded into one's arithmetical language in the usual way.

$\mathrm{Desq}(i, \alpha)$	$=$	the i:th member of the sequence α
$m(g)$	$=$	the m in (4.1) as a function of g
$n(g)$	$=$	the n in (4.1) as a function of g

If $\beta = \langle z_1, z_2, \ldots, z_{m(g)} \rangle$, then $\mathrm{rel}(i, j, \beta, g) = z_{ij}$ where i, j are as in (4.2).

In order to prove that these relations and functions are actually representable in elementary arithmetic we may use the tactic of "proofs by Church's thesis" explained above.

By means of these auxiliary concepts, the desired truth can be defined. It has the initial quantifier prefix

(4.4) $(\exists f)(\forall \alpha)(\forall i)(\forall j)$

The rest is a conditional with the following antecedent:

(4.5) $\mathrm{Seq}(\alpha, m(g))\ \&\ (1 \leq i \leq m(g))\ \&\ (1 \leq j \leq n(g))$

The consequent is a conjunction of two conjuncts. The first is the crucial one. Its function is to state that the truth predicate applies to the Gödel number of a certain quantifier-free and slash-free sentence. This sentence is obtained by substitution from the formula

(4.6) $S[x_1, x_2, \ldots, x_{m(g)}, y_1, y_2, \ldots, y_{n(g)}]$

The Gödel number of (4.6) is by Church's thesis a function of the Gödel number of (4.1) representable in our language. Likewise, so is the Gödel number of the result of all the following substitutions in (4.6) (cf. Mendelson 1987, p. 152):

(4.7) $r(\mathrm{desq}(i, \alpha))$ for x_i
(4.8) $r(f(j, \alpha))$ for y_j

By the T-theorem, the truth predicate applies to the Gödel number of the resulting sentence if and only if the following sentence is true:

$$
\begin{aligned}
(4.9)\quad & S[\mathrm{desq}(1,\alpha), \mathrm{desq}(2,\alpha), \ldots, \mathrm{desq}(m,\alpha)),\\
& \quad f(1, \mathrm{desq}(1,\alpha), \mathrm{desq}(2,\alpha), \ldots, \mathrm{desq}(m,\alpha)),\\
& \quad f(2, \mathrm{desq}(1,\alpha), \mathrm{desq}(2,\alpha), \ldots, \mathrm{desq}(m,\alpha)),\\
& \quad \ldots\\
& \quad f(n, \mathrm{desq}(1,\alpha), \mathrm{desq}(2,\alpha), \ldots, \mathrm{desq}(m,\alpha))]
\end{aligned}
$$

A formula of this form is certainly expressible in our underlying arithmetical language.

The second conjunct is needed for the sole purpose of making sure that some of the functions $f(j,\alpha)$ are independent of some of the variables (members of the sequence α) that they apparently depend on. This can be accomplished by the following clause:

$$
\begin{aligned}
(4.10)\quad & (\forall\beta)(\forall\delta)((\mathrm{Seq}(\beta, m(g)) \ \&\ \mathrm{Seq}(\delta, m(g)))\\
& \&\ (\forall k)(\mathrm{rel}(k, j, g, \beta) = \mathrm{rel}(k, j, g, \delta))) \supset (f(j,\beta) = f(j,\delta)))
\end{aligned}
$$

This completes the explanation of the truth predicate. Unlike e.g., Tarski-type truth definitions, this one is a direct one. It does not relate recursively the truth and satisfaction properties of the simplest types of sentences to the semantical predicates of more complex ones step by step. Ours gives the truth condition of a sentence like (4.1) in one fell swoop as a function of its Gödel number.

Like our original arithmetical truth predicate (2.1), the expression (4.7) is a $\Sigma_1^1\mathcal{M}$ formula. It is therefore (see Walkoe, 1970) expressible in the corresponding IF first-order language, that is, the language for which it is a truth predicate is defined. Before discussing how this can be done, certain general comments are in order.

So far, only arithmetical IF first-order languages have been considered. The treatment can nevertheless be extended to arbitrary first-order languages (as long as they contain elementary arithmetic) and to their arbitrary models. What we have to assume is the following:

> There is representable in the language a function $r(x)$ which maps each member of the domain of individuals $\mathrm{do}(M)$ of the model in question on the individual representing it in the Gödel-type numbering.

It follows that the inverse of $r(x)$ is also representable in the language in question. If this assumption is made, the line of thought pursued goes through without any changes. From the remarks on (4.9) earlier in this section it is seen that all the representatives of particular members of $\mathrm{do}(M)$ are integrated out of the final truth predicate. This truth predicate is therefore independent of any particular model of the theory in question. It depends only on the language used by the theory.

§5. Repercussions of the truth predicate. What is the payoff of directly defined truth predicates? The most straightforward part of the answer is that such a truth predicate is apt to bring out the ways in which the truth condition of a sentence depends on its Gödel number—and the ways it does not depend on that number. Here a general comment on the dependence behavior of complex predicates is in order. In IF first-order logic, the dependencies and independencies of quantifiers and propositional connectives on other quantifiers are studied. However, there is a kind of dependence and independence of the quantifiers occurring in a complex predicate $P[x]$ that has not been examined in the earlier literature. It is the possible independence of such a quantifier, say $(\exists y)$, on the free argument x. The independence of $(\exists y)$ on x can be expressed by writing it $(\exists y/x)$. Such an independence can make a big difference. As an example, one may consider the two predicates of real numbers

$$(\forall y)(A(y) \supset (\exists z)(A(z) \,\&\, (|y - x| = |x - z|))$$
$$(\forall y)(A(y) \supset (\exists z/x)(A(z) \,\&\, (|y - x| = |x - z|))$$

These two obviously impose entirely different conditions on x (and A).

As the reader can easily ascertain, independence of an argument of a complex predicate operates by and large just as independence of a sentence-initial universal quantifier would behave if the argument in question were bound to it.

From this it follows that in the truth predicate defined in the preceding section we have to heed the dependencies and independencies of its quantifiers on the Gödel number g, especially when this truth predicate is converted into a IF first-order form.

One question that arises in connection with this conversion is when the conversion can be continued all the way to an ordinary (slash-free) first-order formula. It is known (and in fact easy to see) that this depends essentially on the argument sets of its Skolem functions. Consider, as an example, (4.1). The argument sets of its Skolem functions are

(5.1) $\{x_{j1}, x_{j2}, \dots, \}$

If these finite sets are linearly ordered by class-inclusion, (4.1) is equivalent with an ordinary first-order formula. If not, formulas like (4.1) in general (independently of the matrix S) do not reduce to ordinary first-order logic.

Hence the reducibility of our truth predicate to ordinary first-order logic depends essentially on the argument sets of its Skolem functions. In these functions, the Gödel number g of (4.1) is handled as if it were a value of an initial universal quantifier.

Now what are these argument sets? A review of Section 4 shows what the answer is. The relevant sets are the following:

(5.2) $\{g\}, \{i, \alpha\}, \{j, \alpha\}, \{j, x_1, x_2, \ldots, x_{m(g)}\}$
$\{k, j, g, \beta\}, \{k, j, g, \alpha\}, \{j, \beta\}, \{j, \alpha\}$

The last four cannot possibly be linearly ordered by class-inclusion. Hence when these four are present, our truth predicate cannot be reduced *ceteris paribus* to ordinary first-order logic. This is trivial, however, for the functions whose argument sets these are were needed to spell out the independence of certain variables y_j from certain variables x_i. The interesting question is what happens when we do not consider these last four cases, in other words, what happens when we are dealing with a sentence of ordinary first-order language.

When the original given sentence (4.1) belongs to ordinary first-order logic, the second conjunct of the truth predicate (see Section 2 above) drops out, and consequently so do the last four argument sets of (5.2). Hence in this case it is especially interesting to ask why our truth predicate cannot be reduced to ordinary first-order logic, as it according to Tarski's (1956) famous result cannot. This question means essentially asking why Tarski's impossibility theorem holds.

A glimpse at the first four argument sets of (5.2) (the last two of which can be considered identical) shows that they are not ordered by class-inclusion. Accordingly, our truth predicate does not normally reduce to ordinary first-order logic, even when it is applied to Gödel numbers of ordinary first-order sentences.

Needless to say, the truth predicate that has been formulated can be expressed in IF first-order logic. If the variable g for the Gödel number of (4.1) is thought of as being bound to a universal quantifier, part of the quantifier prefix of the resulting sentence can be expressed as follows:

(5.3) $(\forall q)(\exists m)(\exists n)(\forall \alpha)(\forall i)(\exists \phi/\forall q)$

Here ϕ is the value of $f(i, \alpha)$. It is crucial to the meaning of the truth predicate that $(\exists \phi/\forall q)$ be independent of $(\forall q)$. As was pointed out earlier in this section, such an independence can obtain and be important even when q is not bound to a universal quantifier.

This situation might seem puzzling. Clearly the truth predicate put forward in Section 3 can be constructed mechanically (recursively) from the Gödel numbers of the sentences to which it is to be applied. Moreover, it is know that all recursive functions can be represented in ordinary elementary first-order number theory by means of ordinary (slash-free) quantifiers. How come, then, that the truth predicate of Section 4 cannot be expressed by means of ordinary first-order logic plus elementary arithmetic?

The answer can be seen from (5.2). What are the argument sets of Skolem functions that are not linearly ordered by class-inclusion? Basically, $\{g\}$ and $\{i, \alpha\}$ plus $\{j, \alpha\}$. Now what is different about these variables? Both g and α are numerical variables. But they are used in a different way. When the syntax of an arithmetical language is represented in itself by means of the techniques of Gödel numbering, numerical variables and other numerical terms can play two different roles. Sometimes they speak realistically of numbers *qua* numbers, as they are used when elementary arithmetic is not yet applied to the study of itself. Sometimes they are used for a different purpose, viz. to represent various expressions of the very same language. Numerical terms of one and the same arithmetical language can serve both purposes in principle and in practice without contradictions and without confusion. In particular, the same quantifiers can range over the same entities (numbers) no matter which purpose they serve. But when two quantifiers play the two different roles, they must naturally be informationally independent of each other. This is simply part and parcel of their being used for different purposes. Moreover these two purposes are not unlike the use and mention distinguished from each other in elementary textbooks of logic.

What we find in the case of direct truth definitions is a case in point. The fact that the argument sets of e.g., the Skolem functions $m(g)$ and $f(j, \alpha)$ are different means that the corresponding existential quantifiers depend on different universal quantifiers. This means that the former existential quantifier must be independent of the universal ones on which the latter depends, and *vice versa*.

Such independence cannot be implemented in ordinary first-order logic. This is the reason why truth definitions for ordinary first-order languages cannot be formulated in those languages themselves. The reason is in effect also the reason why Tarski's celebrated impossibility theorem holds. The deep reason is not any intrinsic limitation of the logic of first-order quantifiers. Nor is it any difficulty about self-reference. It is the need in self-applied arithmetical languages of segregating from each other mentioning (talking about) their various expressions and using them for the purposes of doing number theory, together with the fact that this segregation can only be carried out adequately by means of informationally independent quantifiers.

This diagnosis of the real reasons for Tarski's undefinability theorem has major philosophical implications. First, it demystifies Tarski's result. What is at issue in it is not any deep limitation of human thought, logic, or even any intrinsic limitation of first-order languages. What is involved is nothing deeper than the wherewithal needed to keep use and mention separate in self-applied languages.

This does not mean that the problem of doing so is trivial. But is it a combinatorial problem rather than a philosophical one. There is in fact a certain partial similarity of logical structure between our truth predicate and some of the best known results in combinatorial theory. For instance, the simple ("abridged") version of Ramsey's theorem has the logical form something like "for any natural number k there is a number n such that for any given symmetrical relation R there is a subset X of n (interpreted as the class $\{0, 1, 2, \ldots, n-1\}$) of cardinality k such that the R holds between all or none of the members of X," in symbols

$$(5.4)\quad (\forall k)(\exists n)(\forall R)(\exists X)(X \subseteq n \;\&\; \|X\| = k \\ \&\; (\forall x)(\forall y)((x \in X \;\&\; y \in X \;\&\; x \neq y) \supset (Rxy \;\&\; Ryx)) \\ \vee(\forall x)(\forall y)((x \in X \;\&\; y \in X \;\&\; x \neq y) \supset (\sim R(xy \;\&\; \sim Ryx))))$$

(See here e.g., Graham et al., 1980, chapter 1.)

As Jaakko Hintikka (1996, chapter 4) has pointed out, here the existential quantifier $(\exists X)$ has to be taken to be independent of $(\forall k)$. But if so, the quantifier prefix of (5.4) is similar to (5.3) and the need for informationally independent quantifiers in the truth predicate should not come as any surprise. This is a further indication of the intrinsic connection between truth-definitions and informational independence. It is the failure of earlier logicians to recognize IF quantifiers that has in my opinion prevented them from pursuing direct truth-definitions of the kind presented.

After this side glance, what else is there to be said of Tarski's impossibility theorem? The analysis offered above of direct truth predicates yields not only a diagnosis; it also yields a cure. All that is needed is a way of keeping quantifiers used for different purposes informationally independent of each other. This way is provided by IF first-order logic. Moreover, doing so does not depend on introducing an ingenious new trick into the repertoire of ordinary first-order logic. It results merely from removing from the received logic of quantifiers certain unnecessary restrictions that the notation developed by Frege, Russell and Whitehead and others imposes on patterns of dependence and independence between several quantifiers. Indeed, the terminology used here does injustice to what is being called independence-friendly first-order logic. This logic is simply the general logic of first-order logic *simpliciter*, and to refer to it as a special kind of first-order logic is therefore misleading. Instead of calling it independence-friendly logic one should rather refer to the so-called ordinary first-order logic as dependence-handicapped logic. (Or would the politically correct term be "dependence-challenged"?)

Hence the attention Tarski's impossibility result has received is in a sense due to a historical accident. It is due to the failure of the pioneers of contemporary logic to formulate their logic of quantifiers so as to be free from dispensable, artificial restrictions.

These observations put into an interesting perspective also the prospects of truth predicates outside first-order logical languages. Since all that is needed for such truth predicates is IF quantifiers and since these are present as soon as the use of our quantifiers is not unnecessarily restricted, the natural expectation is that a truth predicate is definable for any language in which quantifiers are used without restrictions. This clearly includes ordinary language, Tarski's "colloquial language," for IF quantifiers can be shown to occur there without restrictions. (Cf. Hintikka, 1974.) Thus what has been found here puts a new complexion not only to the question of the definability of truth in formal languages (for the same language), but likewise to the philosophically central question of the definability of truth in our ordinary working language for that very language—or at least for significant fragments thereof. As is well known, Tarski maintained that the notion of truth cannot be used explicitly and consistently in the colloquial language or any universal language of science. The results obtained here suggest a contrary view. For any language in which the interplay of quantifiers is not needlessly restricted, the definability of truth in the same language is the obvious null hypothesis.

Thus the entire complex of problems concerning the definability of truth thus goes back to the melting pot. What the full story is that is likely to emerge will not be discussed further here. I will not discuss, either, the intriguing historical problem as to why Tarski and indeed the entire philosophical community missed the possibility of self-applied truth definitions on the first-order level and more generally speaking missed to potentialities of direct truth predicates. These questions are discussed in another paper (Hintikka and Sandu, forthcoming).

Our insights thus help to demystify Tarski's famous impossibility theorem. It has nothing to do with the complexity of the languages involved. It does not indicate any deep limitations of human knowledge, as some people have claimed.

These observations can be used to throw some critical light on the so-called disquotational theories of truth. The temptation to embrace such a theory is essentially the same as the fallacious temptation of saying that we can use (4.3) as a truth-condition of (4.1). What is wrong or perhaps rather oversimplified in the usual disquotational analyses of truth is, first, that their proponents do not take seriously Tarski's injunction to consider a quoted expression merely as a complex term whose structure has to be calculated from it (in principle) rather than directly displayed. The second mistake here is a failure to appreciate the combinatorial process involved in the passage from g to the truth condition (4.7) and in particular to underestimate its logical structure. When it comes to serious truth definitions, disquotational accounts do not have an edge over other approaches to truth.

§6. Truth definitions in axiomatic set theory. A truth predicate like the one characterized in Section 4 can be defined for an ordinary or IF first-order theory as soon as it contains elementary arithmetic. The usual first-order axiomatic set theories are cases in point. The resulting truth predicate will be in the relevant IF first-order language. What this implies is that, when applied to axiomatic set theory the resulting truth predicate is not itself set-theoretical. Moreover, by Tarski's result, there is no way of expressing the truth definition in the language of set theory. Hence set-theoretical truth cannot be defined by means of first-order axiomatic set theory.

There is little new in this observation. It is a simple consequence of the fact that the usual axiomatic set theories are first-order theories. The more interesting results are found only by pursuing the same line of thought further.

It might seem that there is a natural way of bringing a self-applied notion of truth into axiomatic set theory as a kind of approximation. The truth predicate defined above is a second order predicate. But if we are to believe Quine, second-order is nothing but "set theory in sheep's clothing." Hence there should be a way of dealing with the notion of truth in set theory, if Quine is right. And there does seem to be a natural way of accommodating the kind of truth predicate we have formulated in set theory. However you look at it, the force of the truth predicate as applied to a sentence S is to assert that the Skolem functions of S exist, just as in our arithmetical language. Now we cannot express such truth definitions directly in axiomatic set theory itself, for they contain either second-order quantifiers or independent (slashed) quantifiers, neither of which occur in axiomatic set theory in its current forms. But, in a perfectly natural sense, we can translate our second-order truth definitions into the language of axiomatic set theory. This certainly must be doable if second-order logic is to be "set theory in sheep's clothing." Indeed, a simple and straightforward translation appears possible. All we need to do is to interpret second-order quantification as quantification over sets of the relevant kind, including functions. Such a translation is possible for any second-order formula, including $\Sigma^1_1\mathcal{M}$ formulas. For instance, let us assume that the following is a second-order formula with $(\exists X)$ (where X is a one-place predicate variable) as its only non-first-order ingredient:

(6.1) $(\exists X)S[X]$

Its obvious set-theoretical translation is

(6.2) $(\exists\alpha)S^*[\alpha]$

where $S^*[\alpha]$ is obtained from $S[X]$ by replacing every atomic formula of the form $X(b)$ or $X(y)$ by $(b \in \alpha)$ or $(y \in \alpha)$ respectively.

Likewise, suppose we have a second-order formula with a one-place initial function quantifier $(\exists f)$ as its only non-first-order ingredient, say

(6.3) $(\exists f)S[f]$.

Its purely set-theoretical translation is of the form

(6.4) $(\exists\alpha)((\alpha \text{ is a one-argument function}) \ \& \ S^*[\alpha])$

Here "α is a one-argument function" is expressible in the language of set theory and $S^*[\alpha]$ is obtained from $S[f]$ in the same way, mutatis mutandis, as in the transition from (6.1) to (6.2). For instance, $f(b) = d$ will be replaced by $\langle b, d\rangle \in \alpha$, in other words $\{\{b\}, \{b, d\}\} \in \alpha$.

Since our truth predicate (4.7) is of the $\Sigma^1_1\mathcal{M}$ second-order form, we can in this way formulate its counterpart in first-order set theory.

Hence there seems to be an eminently natural way of defining truth for axiomatic set theory in set theory itself.

§7. Does the set-theoretical truth-definition work? Thus it might seem that we can, pace our earlier conjectural pronouncements to the contrary, formulate an eminently natural truth-definition for axiomatic set theory in set theory itself. What more can anyone ask? Yet something is clearly still fishy here. For Tarski's impossibility theorem tells us that a fully adequate truth definition for an ordinary first-order theory like axiomatic set theory cannot be given in that theory itself.

But what is meant by "adequate" here? Tarski's (1956, pp. 187–188) requirement uses his famous T-schema or "convention T." It amounts to requiring that sentences of the following form be all provable on the basis of an adequate definition of truth:

(7.1) $\quad \mathrm{True}(\Pi) \leftrightarrow S$

where Π is a quote or structural description of the sentence to be substituted for "S." In our case, Π is the Gödel number $g(S)$ of S.

A weaker form of Tarski's requirement is that all the sentences of the form (7.1) be true, whether or not they are provable.

In our case, the nontrivial halves of Tarski's T-schema are the conditionals that say that the Skolem functions of true sentences exist. An instance of these is the sentence

(7.2) $$(\forall x)(\exists y)S[x, y] \supset (\exists f)(f \text{ is a one-argument function } \& \ (\forall x)S[x, f(x)])$$

This can obviously be expressed in the language of axiomatic set theory, if S is expressed in such a language.

From the failure of the set-theoretical truth predicate it therefore follows that not all conditionals asserting the existence of Skolem functions of a true sentence are true. In other words, in any model of axiomatic set theory there are true sentences whose Skolem functions do not exist in that model in the usual sense of set existence.

This is a telling symptom of the unnaturalness of axiomatic set theory. For one thing, it surely must be the case that the Skolem functions of true set-theoretical sentences must exist if axiomatic set theory is to be able to incorporate its own model theory in any realistic sense. For what Skolem functions do is to show the existence of the witness individuals that are needed so to speak to attest to the truth of those sentences they are Skolem functions of. It is patently absurd to assert a sentence of axiomatic set theory (or of any theory, for that matter) without being prepared to assert the existence of its Skolem functions. For instance, it is plainly impossible to assert an existentially generalized sentence, say

$$(7.3) \quad (\exists\alpha)S[\alpha]$$

without being committed to maintaining that there is some particular individual β of which one can assert

$$(7.4) \quad S[\beta]$$

This idea is codified in the very rule of existential instantiation of first-order logic. It is equally impossible to assert a sentence containing dependent existential quantifiers, say

$$(7.5) \quad (\forall\alpha)(\exists\beta)S[\alpha, \beta]$$

without being committed to maintaining that there is always a way of picking out the requisite witness individual β depending on the value of α, in other words, without also being prepared to assert

$$(7.6) \quad (\exists f)(\forall\alpha)(f \text{ is a function } \& \ S[\alpha, f(\alpha)])$$

And the same holds of course for the relationship of any sentence to the sentence asserting the existence of its Skolem functions.

The simplest example of a sentence whose Skolem function which does not exist in the models of axiomatic set theory is undoubtedly $(\forall x)(\exists y)(x = y)$. Its only Skolem function is of course the identity map of the whole universe, which does not exist in the usual axiomatic set theories. This example was mentioned to me by Professor A. Blass, who describes the situation by saying that here "the translation from second-order logic to set theory replaces the notion of 'universe-sized function' with that of 'set-sized function"' which is enough to change true statements to false ones.

It is nevertheless important to realize that the true set-theoretical sentences without Skolem functions do not always depend on the inexistence of large sets (including functions).

At this point a hard-boiled set theorist might very well object to my procedure in setting up a truth-definition for axiomatic set theory in axiomatic set theory itself. That definition is to capture the notion of truth, it must be presupposed that Skolem functions exist in a model of axiomatic set theory for each and every sentence true in that model. But it is merely a fact of a set theorist's life, it may be alleged, that many sets we would like to see existing in the models of axiomatic set theory just cannot be forced to exist by the current axioms of set theory nor presumably by any finite (first-order) axiom system of set theory. Gödel showed that even elementary arithmetic is inevitably incomplete. *A fortiori*, it is unrealistic to require that this or that set or class of sets that intuitively looks like having a good prima facie claim to existence is not forced to do so by the actual axioms of set theory.

This is true enough in general, but the sets (including functions) which have to exist for my set-theoretical self-applied truth-definition to do its job are not just any old functions. They are Skolem functions of set-theoretical sentences S. And what Skolem functions of S do is to provide us with the witness individuals which are needed to vouchsafe the truth of S. One can have a consistent set theory without the existence of some such functions, but at a cost. That cost is giving up any hope of formulating a model theory of set theory within the selfsame theory. Whatever your conception of truth is or may be, you cannot develop a sensible model theory if you do not have available to you the individuals that make your sentences true. And what Skolem functions do is precisely to furnish you with those witness individuals.

But the shortcoming we are dealing with here is much worse than the general failure of set theory to force certain sets to exist in *all* of its models that we are intuitively inclined to assume to exist. The failure of our truth predicate to be adequate means that in *every* model of axiomatic set theory there are sentences that are false even though they should clearly be true according to our intuitions. And these intuitions are not set theoretical. They do not deal for instance with the existence of very large sets or with any requirement that the models we want to accept be standard ones. They are not the exclusive property of classical mathematicians and logicians, either. On the contrary, constructivistic logicians are likely to require, not merely the existence of any old Skolem functions of all true sentences, but the existence of constructive Skolem functions. These intuitions are not set theoretical, but in a perfectly good sense of the word combinatorial.

§8. Skolem functions and logical truth. This point is worth elaborating further. It may be argued that the existence of the Skolem functions of all true sentences should be thought of as a logical or combinatorial truth rather than set theoretical truth. Indeed, formulated in the framework of second-order logic, it is a logical truth. For in the second-order logic statements asserting the existence of Skolem functions of true sentences are logical truths. The simplest case in point is constituted by the schema

(8.1) $(\forall x)(\exists y)S[x, y] \supset (\exists f)(\forall x)S[x, f(x)]$

where S is free of quantifiers and disjunctions. Hence some set-theoretical counterparts of second-order logical truths will be false in the models of axiomatic set theory. (More will be said about this matter in Section 10 below.)

But we do not have to resort to second-order logic in order to argue our point. For (8.1) is equivalent to a $\Sigma^1_1\mathcal{M}$ sentence, and so are all the second-order sentences asserting the existence of Skolem functions for true sentences. Now it is known that all $\Sigma^1_1\mathcal{M}$ sentences can be translated into the corresponding IF *first-order* language. (By a "corresponding language," we mean a language with the same nonlogical primitives.) This translation turns statements like (8.1) into *first-order* logical truths, that is, logical truths in the same sense and on the basis of the same rules of truth as the logical truths of ordinary first-order logic.

When these logical truths are expressed in set-theoretical jargon, they turn out not only not to be all logical truths. Inevitably some of them will be false. In other words, in each model of axiomatic set theory some formulas that should be true for purely logical reasons are actually false. In other words, in each model of axiomatic set theory some formulas that should be true for purely logical reasons are actually false. This is clearly a serious indictment of axiomatic set theory.

In a sense, we do not even have to go to IF first-order logic in order to appreciate the fact that the Skolem functions of true sentences must exist for logical reasons. The same point can be illustrated by reference to the rules of inference of ordinary first-order logic. When we move from ordinary first-order logic to IF first-order logic, it turns out that the usual rules of inference of first-order logic have to be generalized.

The usual rule of existential instantiation takes us from a formula of the form

(8.2) $(\exists x)S[x]$

to one of the form

(8.3) $S[\beta]$

where β is a new individual constant.

It can be argued that a modification of those rules would bring out more directly the crucial relations of dependence and independence between different quantifiers than the usual set of rules while preserving the first-order character of our logic. A generalized rule of existential instantiation that could do so would take us from a sentence of the form

(8.4) $S_1[(\forall x)S_2[x]]$

where S_1 is in the negation normal form, to

(8.5) $S_1[S_2[f(y_1, y_2, \dots)]]$

where f is a new function constant ("dummy functions" as some logicians might say) and $(\forall y_1)$, $(\forall y_2)$, $\dots$ are all the universal quantifiers of S_1 within the scope of which $(\exists x)$ occurs in S_1. This rule will be called the rule of functional instantiation. It is more general than the received rule of existential instantiation in that it brings out the dependencies and independencies of different quantifiers of each other more fully than the usual rule. At the same time, the new inference rule is obviously valid and otherwise unobjectionable. It is for instance a purely first-order rule. No higher-order quantifiers are used in it. It is in ordinary first-order logic conservative extension of the rules: Whatever is provably valid with the help of the new rule (i.e., with the help of new function symbols) is provable by the old set of rules. The old rule of existential instantiation can be considered a special case of the new rule, viz. the case in which we do not have to consider which quantifier depends on which.

Admittedly, one does not necessarily know what the instantiation function is that the new function symbol (such as f in (8.5)) stands for. But this is no objection to the new rule. If it were, it would apply with an equal force to the simple old rule of existential instantiation which takes us from (8.2) to (8.3). For the very purpose of the received rule of existential instantiation can be said to be to enable us to argue about (and by reference to) individuals which are known to exist but of which we do not know what or who they are. Likewise, the generalized rule enables us to argue about (and by means of) functions that are known to exist but are not known to us.

The naturalness, not to say the unavoidability, of the generalized rule is also shown by the fact that when we move from ordinary first-order logic to independence-friendly (IF) first-order logic, the new rule is no longer conservative. It is needed to prove (and/or disprove) formulas that otherwise could not be proved (and/or disproved). The reason is clear. An existential quantifier $(\exists x/\forall y)$ independent of a universal quantifier $(\forall y)$ is typically stronger than the corresponding dependent existential quantifier $(\exists x)$. It says that a certain choice can always be made on the basis of less information than in the case of a dependent existential quantifier, which is a stronger claim than the one the dependent existential quantifier codifies. This extra strength does not come into play in the usual rules of logical inference where an existential quantifier is always treated as if it depended on all universal quantifiers within whose (syntactical) scope it occurs.

But even though the new rule is indispensable only in IF first-order logic, not in ordinary one, there is nothing in it that depends on the presence of IF quantifiers. It is a valid rule already in ordinary first-order logic, and the reason why it is needed in IF first-order logic is that in it subtler patterns of independence and dependence among quantifiers are allowed than in ordinary first-order logic.

To return to axiomatic set theory, the rule of functional instantiation can naturally be used in it. The only modification is that in the output of the rule the logical type of the instantiating function constant has to be indicated by a separate conjunct. For instance, we can now move from

(8.6) $(\forall\alpha)(\exists\beta)S[\alpha,\beta]$

to

(8.7) $(\exists f)(\forall\alpha)(S[\alpha, f(\alpha)]\ \&\ (f \text{ is a one-place function}))$

where the second conjunction can be expressed in the language of set theory.

What is remarkable here is that the new rule of inference (the rule of functional instantiation, as thus applied in axiomatic set theory) would enable us to prove all the conditionals where the antecedent is a closed sentence S of axiomatic set theory and where the consequent asserts the existence of the Skolem functions for S. But is was seen that these conditionals cannot all be true in any one model of axiomatic set theory. Hence the extended rule of existential instantiation (functional instantiation) which was seen not only as a valid purely logical rule of inference but indispensable in IF first-order logic is seen to yield wrong results when formulated in axiomatic set theory.

§**9. Conclusions.** But what is the upshot of the line of thought we have indulged in? Anyone can formulate a set-theoretical predicate which is a set-theoretical translation of the S truth predicate we have indicated how to formulate, and call it a truth predicate. But is anyone any wiser? A predicate has a hope of being a truth predicate only if it satisfies Tarski's T-schema. (Tarski 1956, pp. 155–156) Now the conditionals we have ostensively been defending are but the nontrivial halves of the different substitution instances of the T-schema. If they are all true in a model of axiomatic set theory, the putative set-theoretical truth predicate is a genuine article. But that is precisely what Tarski's famous theorem shows to be impossible. Hence, for all the naturalness of the set-theoretical truth predicate defined above, it fails to do the job. It fails because it flunks the T-schema test. In any given model M of axiomatic set theory, there are sentences true in M but whose Skolem functions do not exist in M.

Since the first and foremost foundational mission of axiomatic set theory was seen to provide for a model theory, including a truth predicate, for all and sundry systems, it fails miserably. It cannot even provide a model theory for itself.

But this failure shows more than merely the failure of axiomatic set theory to serve as its own model theory. It casts a shadow on axiomatic set theory in a more radical way. This shadow is cast by the naturalness of the requirements that any satisfactory axiomatization of set theory must satisfy. We argued in Section 8 above for the naturalness of the critical conditionals (instances of the T-schema). The conditionals do not merely look like plausible candidates for being considered set-theoretical truths. In their non-set-theoretical incarnations, they are logical (combinatorial) truths, in that they are logical truths of IF first-order logic. Axiomatic set theory is thus not only incomplete in that it does not allow for a truth predicate to be defined for it in the same language. In each model of axiomatic set theory there are true sentences which are in a natural sense of the expression combinatorially false.

This perverse character of axiomatic set theory is illustrated by our further results. The truth of combinatorially false sentences (in an intuitive sense of "combinatorially false") of course does not mean that axiomatic set theory is inconsistent. But what was found shows that axiomatic set theory becomes formally inconsistent if the rules of inference of the underlying logic (which is merely our ordinary first-order logic) are changed—or, rather, when the modified rules of inference are expressed in the language of axiomatic set theory. In a sense, not only is it impossible to develop a model theory of axiomatic set theory in that theory itself. In that theory, we cannot even formulate those deductive rules of first-order logic that are the most natural ones model-theoretically.

Furthermore, since combinatorial truths which fail in axiomatic set theory are truths of IF first-order logic, it is literally true that in any model of axiomatic set theory certain *logical truths* are false. Only the fact that the logic involved in these negations of logical truths goes beyond the ordinary first-order logic employed in axiomatic set theory saves this theory from outright contradictions.

It may be in order to pause here and to contemplate for a moment the consequences of our results. We have discovered that there are intuitively (combinatorially) false sentences that hold in the models of axiomatic set theory. These false sentences even include in a sense logical truths. However, this does not mean that this theory is inconsistent. It does not mean that it is not worth studying as one mathematical theory among many. This theory can also have important application to other mathematical theories, such as measure theory or theory of real functions. What the discoveries reported here do, however, is to deprive axiomatic set theory most of its foundational interest. For example, consider questions of independence. Since there are false sentences in all the models of axiomatic set theory, the question whether or not some given conjecture, such as the continuum hypothesis or its negation, is implied by the axioms of set theory has no intrinsic connection with the truth of the hypothesis in question. *A fortiori*, results like Gödel's (1940) or Paul Cohen's (1966) independence theorems have no direct relevance to the question of the truth or falsity of the continuum hypothesis. The question of the actual truth of the continuum hypothesis and of similar conjectures must be approached by means other than ordinary first-order axiomatic set theory. Furthermore, there is no point in looking for new axioms of set theory to be added to the old ones, for they can only exacerbate the situation.

It is perhaps objected here that this conclusion depends on thinking of axiomatic set theory as being second-order logic in wolf's clothing, that is to say, depends on interpreting theorems of axiomatic set theory as attempts to capture truths of second-order logic on its standard interpretation. The answer is that the very meaning of propositions like the axiom of choice, the continuum hypothesis etc. in axiomatic set theory depends on the same way of looking at set theory. If you try to dispense with that way of thinking, you cannot any longer claim to be speaking of the axiom of choice, of the continuum hypothesis etc. in your set theory.

More generally speaking, what we have here is the ultimate reason for the failure of axiomatic set theory to serve as a generic model theory of various mathematical or scientific systems. From an authentic model-theoretical viewpoint, axiomatic set theory is bound to give some false results.

§**10. A perspective on the axiom of choice.** There is a way of looking at our line of thought which might at first sight seem to suggest a way of criticizing it. It might be suggested that we are tacitly relying on the axiom of choice or perhaps a generalized form of the axiom. And the axiom of choice might not be, as its history shows, beyond reasonable (or unreasonable) doubts. (For such doubts, see Moore (1982).) For instance, the game-theoretical truth-condition of sentences of the form

(10.1) $(\forall x)(\exists y)S[x, y]$

viz.

(10.2) $(\exists f)(\forall x)S[x, y]$

is implied by the former only in virtue of the axiom of choice. More generally, the same goes for the Skolemization Lemma, which can almost be considered as a generalized form of the axiom of choice. Hence our entire approach might seem to be subject to all the same doubts and criticisms as the axiom of choice.

There is a point to such criticisms, but they suffer—to put the matter bluntly—of the same prejudice which has led people to brand higher-order logic as set theory in disguise. The view that emerges from our argument is, rather, that the axiom of choice is a valid logical principle of second-order logic which is only imperfectly codified in its formulation as an axiom of set theory.

It has been detrimental to the correct understanding of the axiom of choice that it has been thought of as a set-theoretical axiom. This is in fact not how some of the keenest minds in the field have thought of this principle. For instance, no less a figure than Hilbert (1922, p. 157) thought of the axiom of choice as a logical principle which he hoped to show to be as obvious as $2 + 2 = 4$. And if we look at such conditionals as

(10.3) $(\forall x)(\exists y)S[x, y] \supset (\exists f)(\forall x)S[x, f(x)]$

we can see the truth of their view. Conditionals like (10.3) are truths of second-order logic, and they receive their importance from the fact that they are logical bridges between first-order and second-order languages. If you do not first believe this, you are invited to translate (10.2) into an IF first-order language. There it becomes a first-order logical truth, quite as genuine one as the logical truths of ordinary first-order logic. As is shown by the

Skolemization Lemma, they express in effect truth conditions of first-order sentences.

Hence the proper conclusion is not that our argument is subject to the same criticisms as the axiom of choice. Rather, our analysis of the truth conditions of first-order sentences constitutes a vindication of the axiom of choice, in the sense of showing that all its instances are in effect logical truths.

Moreover, the force of the logical truths that are manifested in the second-order form of the axiom of choice is not exhausted by the usual set-theoretical formulations of the axiom. The axiom of choice asserts the existence of a function which picks out a unique element from each set of a nonempty *set* of *sets*. But in a second-order conditional like

(10.4) $(\forall x)(\exists y)S[x, y] \supset (\exists f)(\forall x)S[x, f(x)]$

the classes

(10.5) $\{y : S[x, y]\}$

for the different x's need not all be sets, nor need the class of all such sets (or classes) be a set. And if not, the logical force of (10.4) is not fully captured by the set-theoretical form of the axiom.

These observations can be turned into a new criticism of axiomatic set theory. Why don't we simply add all these instances of a generalized axiom of choice to the received axioms of set theory? The answer is that we can do so only in the pain of inconsistency. For one way of restating the conclusions of the earlier part of our paper is that the totality of all the instances of the schema (10.4) in their set-theoretical formulation is incompatible with the axioms of set theory. This is objectionable already from the perspective of classical set theory. The reason is that whatever theoretical motivation there is for assuming the axiom of choice applies directly for the defense of schema (10.3) without any restrictions. The conclusion is, therefore, that the full force of the intuitions backing the axiom of choice cannot be implemented in axiomatic set theory. In short, in this sense the axiom of choice is incompatible in its fully general form with the usual axiomatic set theories.

Insofar as the mission of axiomatic set theory includes implementing the intuitive ideas behind the axiom of choice, it is bound to fail in this direction, too.

One can say more than this, however. We have argued that all the different instances of generalized axiom of choice are *first-order* logical truths and hence in a sense combinatorial rather than set-theoretical truths. For what they assert is the existence of Skolem functions for a sentence S, given this sentence. And that existence is the very truth-condition of S. Therefore all these conditionals must be thought of as logical truths. And a translation into IF first-order logic accordingly turns them into literal first-order logical truths.

Hence the failure of the Skolem functions of some true sentences to exist in the models of axiomatic set theory means that there are some logical truths that are false in those models. Can there be a worse way for axiomatic set theory to fail in its mission?

§11. Is Brouwer our unwitting ally? So far, only the shortcomings of axiomatic set theory have been discussed. In order to clarify the basis of the criticism that has been offered, it is in order to show what consequences the same line of thought are in other parts of the foundations of mathematics. What has been shown is that if a truth predicate is formulated in set theory, we end up with sentences which are true but whose Skolem functions do not exist. Now somebody might feel like offering here for axiomatic set theory the classical lame excuse: others are wrong, too. What this excuse points to is that a similar conundrum arises outside set theory, too. For instance, we have seen that a truth predicate can be defined for the $\Sigma^1_1\mathcal{M}$ fragment of second-order logic in the same fragment. But that fragment is not closed with respect to ordinary contradictory negation. When it is extended by allowing sentence-initial contradictory negation, we obtain a language which is consistent but which exhibits the same flaw—or at least the same peculiarity—as the language of axiomatic set theory. It contains sentences which do not imply the existence of their own Skolem functions. They can be true even though the corresponding Skolemized sentence is false.

An example is provided by an attempt to refute the possibility of a self-applied truth predicate for IF first-order arithmetical languages by means of a liar paradox type argument. (See here and in the following Hintikka, 1996, ch. 7.) If $T[n]$ is the truth predicate, we can apply the diagonal lemma to

(11.1) $\sim T[x]$

where $\sim$ is the strong (dual) negation and obtain a number g which is the Gödel number of

(11.2) $\sim T[\mathbf{g}]$

where $\mathbf{g}$ is the numeral representing g. Now it is easy to see that (11.2) is neither true nor false, which is perfectly possible in IF first-order logic. Thus we cannot reconstruct the liar paradox by means of the strong (dual) negation $\sim$. But of course we can introduce also the usual contradictory negation $\neg$ into our first-order logic, can't we? Yes, we can, but it turns out that it cannot be given independent semantical rules for and that it can therefore be prefaced only to sentences (closed formulas). Hence we cannot produce a contradiction by means of it, either. For if $\neg$ is the contradictory negation, $\neg T[x]$ is ill-formed, for $\neg$ cannot be prefixed to open formulas.

Hence there is no number h which would be the Gödel number of $\neg T[\mathbf{h}]$. But we can write

(11.3) $\neg \sim T[\mathbf{g}]$

This sentence says, roughly, "It is not true that I am false," which is true. But there cannot exist Skolem functions for (11.3), for their existence would show that (11.2) is true after all. Hence (11.3) is true but without Skolem functions.

Since the negation of (11.3) fails to be true, Skolem functions do not exist for it, either. In brief, (11.3) is an example of a first-order sentence for whose Skolem functions and the Skolem functions of whose negation fail to exist.

Thus we can find a formula S of extended IF first-order logic such that there exist no Skolem functions either for S or for $\sim S$. And its second-order translation is a perfectly classical sentence of $\Pi^1_1\mathcal{M}$ logic.

This reveals an interesting perspective on intuitionistic ideas. Jaakko Hintikka (1996, ch. 11) has suggested that one crucial turn taken by the original intuitionists was to think in terms of our knowledge of mathematical truths rather than those truths plain and simple. Even more importantly, the intuitionists were focusing on our *knowledge of mathematical objects* of different logical types, especially on the objects we need to know mathematical truths, rather than our *knowledge of mathematical facts or truths*. Skolem functions can be thought of as such mathematical objects in search for a knower. Skolem functions of a first-order sentence are objects that have to be known in order to know mathematical truths, for they produce the witness individuals that bring out the truth of the sentence in question. For instance, what a Skolem function f of

(11.4) $(\forall x)(\exists y)S[x, y]$

does is to produce the witness individual $f(x)$ that for each x verifies

(11.5) $S[x, f(x)]$

Our knowledge of mathematical objects is thus important for intuitionists because it is our means of coming to know mathematical truths. Intuitionists claim that we must have the wherewithal of actually knowing the mathematical propositions we assert. A characteristic twist given by the intuitionists to this requirement is that we must know the mathematical objects that serve to verify those propositions. Those objects are obviously in the case of first-order sentences the functions that produce the witness individuals, in other words, Skolem functions. Jaakko Hintikka (1996, chapters 10–11) has proposed implementing at least experimentally the knowability requirement by requiring that those critical functions be required to be knowable functions in some sense or other, for instance computable (recursive) functions. Such requirements do in fact give rise to interesting ways of looking at mathematical propositions and mathematical theories. For instance, in this way we can obtain a complete axiomatization of elementary arithmetic on the first-order level. These ways of implementing the intuitions of the intuitionists are interesting and eminently worth being examined in detail and developed further.

What we have found here is a far deeper source of disquiet for any thinker who is concerned with the knowability of mathematical objects and propositions. What we have found is that in the presence of contradictory negation there are (both in axiomatic set theory and in extended IF first-order logic) true sentences whose witness functions not only are unknowable, but do not even exist. (Nonexistence is the strongest possible reason for unknowability, one is tempted to say here.) Moreover, it is the introduction of contradictory (classical) negation that gives rise to true sentences without Skolem functions. For there are no such sentences in plain IF first-order logic. We have to ascend to the *extended* IF logic in order to be able to find them.

Now a Brouwer might use the presence of such unverifiable sentences in extended IF languages as a way of criticizing them, just as we criticized axiomatic set theories. And the reason why extended IF first-order logic nevertheless avoids contradictions and all other technical problems is that we have followed Brouwer's advice and reformed our logic, among other things by giving up the law of excluded middle for the only negation (the dual or strong negation) for which we can give real semantical rules. Hence our criticism of axiomatic set theory has an interesting but hidden similarity with Brouwer's criticism of classical mathematics.

Of course, Brouwer did not consider in so many words truths of extended IF logic, for the good reason that no such logic was in existence in his time. But, as Jaakko Hintikka (1996, ch. 4) has pointed out, there are concepts and propositions in more or less ordinary mathematics which tacitly involve independent quantifiers. If we apply classical logic to them, we very easily end up at the very least with sentences whose Skolem functions do not exist, and the Skolem functions of whose negations do not exist, either. Thus we have located a true gem of a problem for historians of mathematics: to find such tacitly independence-friendly propositions in Brouwer's actual mathematical work.

Of course technically we can escape with our classical logic unscathed if we simply move to second-order logic. But interpretationally that would mean, not just asserting mathematical propositions, but inquiring into their witness functions (Skolem functions). In the eyes of someone like Brouwer, this could easily look like talking about mathematical truths rather than asserting and proving them.

Thus we can see, not only that there is a connection between our criticism of axiomatic set theory and Brouwer's intuitionistic ideas, but also that

Brouwer is right in that such criticisms motivate a critical look at the very logic we all use in mathematics.

For instance, in a Brouwerian perspective we must be wary of extended IF logic, just because it creates true sentences without Skolem functions. In his (1996, ch. 9), Jaakko Hintikka emphasized that practically all of normal mathematics can in principle be formulated in the framework of IF first-order logic. Now we have found some reason for the inverse conclusion, viz. that there are good reasons in principle not to venture beyond the (unextended) IF first-order logic.

§12. The confusions of intuitionistic logic. Thus the predicament that the original intuitionists more or less explicitly pointed out is not due to any problematic set-theoretical assumptions. It arises in extended IF first-order arithmetic. It requires—here the intuitionists are entirely right—a critical look at the very logic used in mathematics. However, this critical look cannot be restricted to classical (contradictory) negation and it does not reduce to blaming the malaise of classical mathematics on classical negation. Tinkering with the formal laws of negation is a cheap thrill. Any justification for changes in the laws of governing negation must be based on an analysis of the underlying semantical (model-theoretical) situation. We have carried out such an analysis, and what it shows is that the real culprit (or hero) here is not negation. Indeed, all the semantical rules for negation remain intact when we move from ordinary first-order logic to IF first-order logic. (Admittedly, some deductive rules may change, but merely because they are parasitic on the semantical rules.) What changes is that the interplay of quantifiers is liberated from unnecessary restrictions.

Unfortunately, these deep issues usually get in discussions of intuitionism enmeshed with all sorts of extraneous issues and frequently confused. For instance, as was pointed out, there is a simple expedient to ensure the existence of the Skolem functions of all true sentences, viz., to use only IF first-order logic. In Hintikka (1996, ch. 9) it has been shown that a great deal of mathematical reasoning can in fact be carried out in this logic. But we do not capture this logic fully even on the deductive level unless we strengthen the rule of existential instantiation so as to allow in effect an instantiation of Skolem functions and not only of existentially bound variables. (This does not mean transcending the first-order level.) Yet in Heyting's formalization of intuitionistic logic there is no such a rule of functional instantiation. Hence this logic can scarcely be said to be a full implementation of the intuitions of the original intuitionists.

This point is connected with the changing attitudes of the intuitionists to the axiom of choice. The original intuitionists took in effect a proposition of the form

(12.1) $(\forall x)(\exists y)S[x, y]$

to make the purely existential claim that there is some way or other way of picking a truth-making value y in $S[x, y]$, given any value of x. From this it of course does not follow that there is a known function that does the job, even if we know that (12.1) is true. But this line of thought presupposes that the quantifiers of (12.1) are interpreted classically. If they are interpreted intuitionistically, (12.1) will say the same as its purported but allegedly invalid conclusion. In other words, if quantifiers are interpreted intuitionistically in a consistent manner, the axiom of choice is intuitionistically valid. This conclusion has in effect been acknowledged by some leading latter-day intuitionists such as Dummett (1977, pp. 52–53) and Martin-Löf (1984, pp. 50–54).

This is the predicament of constructivistic logicians more generally. They are trying to capture the logical behavior of constructivistically interpreted quantifiers. But if so, they cannot maintain that the laws that are to be rejected *when quantifiers are interpreted classically* must also be rejected in their formal system.

This applies in a sense even to the law of excluded middle. Classically, the sentence

(12.2) $(\forall x)(A[x] \vee \sim A[x])$

is equivalent with the corresponding "Skolemized" sentence

(12.3) $(\exists f)(\forall x)((A[x] \,\&\, f(x) = 0) \vee (\sim A[x] \,\&\, f(x) \neq 0))$

But if the function variable f is in (12.3) restricted to recursive or otherwise knowable functions, (12.2) no longer entails (12.3). This can be thought of as capturing the gist of the motivation of the intuitionistic rejection of the law of excluded middle. But an intuitionistic logician cannot formulate his or her point by speaking of the failure of (12.2) to imply (12.3), because an intuitionist will have to interpret the disjunction in (12.2) intuitionistically in his or her own logic. And this would restore the equivalence between (12.2) and (12.3).

Thus the very presuppositions of the so-called intuitionistic logic have to be examined critically. In view of all this confusion, it is sorely tempting to apply to intuitionism the same line that Karl Kraus used of psycho-analysis: it is the illness it pretends to be a cure of.

On the conceptual level, one crucial distinction that is missing from much of the discussion concerning intuitionism is the distinction between the different senses of verification. (Cf. here Hintikka, 1996, ch. 2.) On the one hand, verifying an interpreted first-order sentence may mean using its Skolem functions actually to calculate (or otherwise to find) the witness individuals that "prove" its truth. On the other hand it may mean finding the Skolem functions that guarantee the success of the former enterprise. These are not two competing accounts of verification. The latter does not represent a pure existence proof as distinguished from a constructive one. If Brouwer thought so, he was a victim of one of the very confusions he was trying to eradicate. The two are conceptually (categorially) different. The former is like successfully defending a claim against a challenge. The latter is like finding a strategy that is guaranteed to enable us to meet all such challenges. The former is like finding a tactic in some particular ballgame against some particular opponent that leads to a win. The latter is like hiring a coach who can find such a ploy in all the games of the season.

This contrast is not identical with the contrast between truth and provability, either, nor with the similar distinction between material and logical truth. Both terms of our contrast pertain to the truth of sentences, not to their provability or their logical truth, but in a different way. In one of them, we are dealing with the moves in a certain game, in the other with strategies in the same game.

The difference is not analogous with the difference between the two kinds of negation, either, even though there is a close relationship. The strong (dual) negation of a sentence S means that there is a unified defense against all challenges to $\sim S$, whereas the weak (contradictory) negation $\neg S$ of S simply denies the existence of such an overall strategy for S. (This leaves it possible for the verifier still to win in some of the plays of the game related to S.) The failure of the intuitionists to distinguish the two negations from each other strongly suggests that they were confused in this matter.

One corollary to our diagnosis of the ills of intuitionists is that they are barking up the wrong tree when they blame the shortcomings of classical logic on infinity. The step from ordinary first-order logic to IF first-order logic is rare among the different ways of strengthening ordinary first-order logic in that it makes a difference already to finite models. It is in fact easy to find examples of sentences which in a suitable finite model are neither true nor false. Hence, whatever problems beset classical logic they are not due an illicit extension of our methods of dealing with finite sets to the realm of the infinite, as Brouwer (1929, pp. 158–159 of the original) claims they are.

The constructivistic validity of the axiom of choice seems to be contradicted by what we find in the actual intuitionistic mathematics. As has been pointed out repeatedly (see Scott 1968, Diaconescu 1975, Goodman and Myhill 1978), one can find counter-examples to the axiom of choice in intuitionistic mathematics.

However, what is at issue here is the fact that intuitionism proper implies something quite different from run-of-the-mill constructivism. In the classical Brouwerian intuitionism, we are not just dealing with mathematical truths, constructivistic or not. As I have argued in Hintikka (1996, ch. 11) what is addressed by the likes of Brouwer is our knowledge of mathematical truths and more importantly of mathematical objects. For this vantage point, the axiom of choice does indeed fail, the reason being that our knowledge of the truth of a sentence of the form $(\forall x)(\exists y)S[x, y]$ does not imply a knowledge of any choice function f that would make true the sentence $(\forall x)S[x, f(x)]$. But then the principle that fails cannot any longer be expressed by the classical conditional

$$(\forall x)(\exists y)S[x, y] \supset (\exists f)(\forall x)S[x, f(x)]$$

Thus what has been said about the vindication of the axiom of choice in the eyes of consistent constructivists must be restricted to constructivists as distinguished from intuitionists. What tends to confuse the issues is the unfortunate tendency of many constructivists to call their position "intuitionistic."

REFERENCES

[1] BROUWER, L.E.J., *Mathematik, Wissenschaft und Sprache*, ***Monatshefte für Mathematik***, vol. 36, (1929), pp. 153–164. Reprinted in Brouwer (1975) pp. 417–428; English translation in Eward (1996), pp. 1170-1185.

[2] ———, ***Collected works***, vol. 1, (A. Heyting, editor) North-Holland, Amsterdam, 1975.

[3] COHEN, PAUL J., ***Set theory and the continuum hypothesis***, W. A. Benjamin, Amsterdam and New York, 1966.

[4] DIACONESCU, R., *Axiom of choice and complementation*, ***Proceedings of the American Mathematical Society***, vol. 51, (1975), pp. 176–178.

[5] DUMMETT, M., ***Elements of intuitionism***, Clarendon Press, Oxford, 1977.

[6] EWARD, W. B., editor, ***From Kant to Hilbert***, Clarendon Press, Oxford, 1996.

[7] FRIEDMAN, M., *Logical truth and analyticity in Carnap's 'Logical syntax of language'*, in ***History and Philosophy of Modern Mathematics*** (W. Aspray and P. Kitcher, editors), Minnesota Studies in the Philosophy of Science, vol. XI, Unuversity of Minnesota Press, Minneapolis, (1988), pp. 82–94.

[8] GÖDEL, K., ***The consistency of the continuum hypothesis***, Annals of Mathematical Studies, vol. 3, Princeton University Press, Princeton, 1940.

[9] GOODMAN, N. D. and MYHILL, J. R., *Choice implies excluded middle*, ***Zeitschrift für Mathematische Logik und Grundlagen der Mathematik***, vol. 24, (1978), p. 461.

[10] GRAHAM, R. L., ROTHSCHILD, B. L, and SPENCER, J. H., ***Ramsey theory***, John Wiley and Sons, New York, 1980.

[11] HENKIN, L., *Completeness in the theory of types*, ***The Journal of Symbolic Logic***, vol. 15, (1960), pp. 81–91.

[12] HEYTING, A., ***Intuitionism: an introduction***, North-Holland, Amsterdam, 1956.

[13] HILBERT, D., *Neubegründung der Mathematik. Erste Mitteilung*, ***Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universitä t***, vol. 1, (1922), pp. 157–177. English translation in Eward (1996), vol. 2, pp. 1115–1134.

[14] HINTIKKA, JAAKKO, ***The principles of mathematics revisited***, Cambridge U.P., 1996.

[15] ———, *What is elementary logic?*, in ***Physics, philosophy and the scientific community***, (K. Gavroglu et al., editors), Kluwer Academic, Dordrecht, (1995), pp. 301–326.

[16] ———, *Quantifiers vs. quantification theory*, ***Linguistic Inquiry***, vol. 5, (1974), pp. 153–177.

[17] HINTIKKA, J. and SANDU, G., *Tarski's guilty secret: compositionality*, forthcoming.

[18] ———, *Game-theoretical semantics*, in ***Handbook of logic and language***, (Johan van Bentham and Alice ter Meulen, editors), Elsevier, Amsterdam, (1997), pp. 361–410.

[19] MARTIN-LÖF, P., ***Intuitionistic type theory***, Bibliopolis, Napoli, 1984.

[20] MENDELSON, E., ***Introduction to mathematical logic***, Third edition, Wadsworth and Brooks/Cole, Monterey, CA, 1987.

[21] MOORE, G. H., ***Zermelo's axiom of choice***, Springer-Verlag, New York, Heidelberg and Berlin, 1982.

[22] ———, *Logic and set theory*, in ***Companion encyclopedia of the history and philosophy of the mathematical sciences***, (Grattan-Guinness et al., editors), vol. 1, Routledge, London and New York, (1994), pp. 635–643.

[23] PUTNAM, H., ***Philosophy of logic***, Harper and Row, New York, 1971.

[24] ROGERS, H., JR., ***Theory of recursive functions and effective computability***, McGraw-Hill, New York, 1967.

[25] SANDU, G., *IF first-order logic and truth-definitions*, ***Journal of Philosophical Logic***, vol. 26, (1997).

[26] SCOTT, D., *Extending the topological interpretation to intuitionistic analysis I*, ***Compositio Mathamatica***, vol. 20, (1968), pp. 194–210.

[27] SUPPES, P., ***Introduction to logic***, Van Norstrand, New York, 1956.

[28] TARSKI, A., ***Collected papers***, vols. 1–5, ed. by S. R. Givant and R. N. McKenzie, Birkhäuser, Basel, 1986.

[29] ———, ***Logic, semantics, metamathematics***, Clarendon Press, Oxford, 1956.

[30] TARSKI, A., *Der Wahrheitsbegriff in den formalisierten Sprachen*, ***Studia Philosophica***, vol. 1, (1935), pp. 261–405. Reprinted in Tarski vol. 2, (1986), pp. 51–198; English translation in Tarski 1956, pp. 152–278.

[31] WALKOE, W. JR., *Finite partially ordered quantification*, ***The Journal of Symbolic Logic***, vol. 35, (1970), pp. 535–555.

DEPARTMENT OF PHILOSOPHY
BOSTON UNIVERSITY
745 COMMONWEALTH AVENUE
BOSTON, MA 02215, USA
E-mail: hintikka@bu.edu