Notre Dame Journal of Formal Logic
Volume 37, Number 1, Winter 1996

Tarski on Logical Consequence

MARIO GÓMEZ-TORRENTE

Abstract This paper examines from a historical perspective Tarski's 1936 essay, "On the concept of logical consequence." I focus on two main aims. The primary aim is to show how Tarski's definition of logical consequence satisfies two desiderata he himself sets forth for it: (1) it must declare logically correct certain formalizations of the ω-rule and (2) it must allow for variation of the individual domain in the test for logical consequence. My arguments provide a refutation of some interpreters of Tarski, and notably John Etchemendy, who have claimed that his definition does not satisfy those desiderata. A secondary aim of the paper is to offer some basic elements for an understanding of Tarski's definition in the historical logico-philosophical context in which it was proposed. Such historical understanding provides useful insights on Tarski's informal ideas on logical consequence and their internal cohesion.

1 Introduction Alfred Tarski was one of the founders of modern logic. His exceptionally deep, rich, far-reaching contributions to all areas of research in the field have been and will be an eminent source of inspiration for all those working in mathematical logic. But Tarski's work has been also very influential in philosophy. In particular, his definition of truth for formalized languages was taken as a paradigm of exact philosophical explication of a common concept and later used as an instrument for diverse philosophical enterprises. The philosophical literature on Tarski's definition of truth and on its applications is now very extensive, and as a result our appreciation of its philosophical virtues and limitations has steadily grown.

The situation is different in the case of another example of exact philosophical explication due to Tarski, and closely related to his definition of truth: his definition of logical consequence. According to this definition, a sentence follows logically from a set of sentences if there is no model of the latter which is not a model of the former. Despite the obvious philosophical importance of the notion of logical consequence, philosophical discussions of Tarski's definition are scarce in the literature. This is not, against what one might at first think, a sign that Tarski's definition is philosophically sterile or uninteresting. Philosophical discussions are for the most part critical,

Received February 3, 1995; revised November 14, 1995

or attempts at refutation, and it is critical discussions of Tarski's definition that are scarce. This scarcity is a sign of the philosophical success of the definition, and an even more important but less obvious sign of this success is the uncritical, tacit use of the definition in many of the most important contemporary pieces of philosophical argumentation.

Among the few critical discussions of Tarski's definition of logical consequence in the philosophical literature, some of the most challenging are contained in several recent publications by Etchemendy ([10], [7], [9]). It is to be hoped that Etchemendy's critical discussion, and the responses that it has provoked, may be the basis for a sustained philosophical exploration of the virtues and limitations of Tarski's definition of logical consequence. But a necessary point of departure for such an exploration must be, undoubtedly, a close interpretative reading of Tarski's writings relevant to the topic. Such a close reading can be reasonably expected to yield not only historically interesting data, but also illumination at the conceptual level, for example, by clarifying how Tarski's definition was thought by him to satisfy the pretheoretic intuitions that motivated it.

In this paper I will discuss and analyze from a historical perspective some of the aspects of Tarski's work that are of direct relevance for a philosophical understanding of his definition of logical consequence. There is virtually no literature examining historically this part of Tarski's work. If we exclude some isolated references in works of a more general nature, the only close examination of Tarski's writings in this area is again a paper by Etchemendy ([8]). In the course of my discussion of Tarski I will have repeated occasion to refer to Etchemendy's historical treatment, and I will argue that it contains some errors of interpretation and misapprehensions.

In Section 2 below I summarize some of the central points of the paper in which Tarski proposed his definition of logical consequence (Tarski [30]). I pay special attention to Tarski's description of his reasons for proposing such a definition; of particular importance here is Tarski's remark that, in view of Gödel's results, all formalized theories of arithmetic of a certain sort are ω-incomplete, and yet the ω-rule is intuitively logically valid. I also pay close attention to the fragments more directly related to the statement of the definition of logical consequence itself. The rest of the paper consists of a series of arguments that show how Tarski's definition fulfilled the pretheoretic requirements he had in mind when he gave it.

Section 3 considers the case of the ω-rule. Etchemendy has pointed out, first, that the ω-rule is not logically valid (according to the Tarskian definition) in first-order arithmetic with the usual choice of logical constants; and, second, that if this is so, either the definition does not meet Tarski's pretheoretic requirements or it must resort to an unusual choice of logical constants (a choice that has counterintuitive consequences). I show that Tarski thought that the ω-rule was logically valid only under certain formalizations of arithmetic (roughly, those in which arithmetical notions can be defined in the logicist fashion), but not in first-order arithmetic; hence the dilemma set forth by Etchemendy disappears. A crucial aim of Tarski (that emerges in both Sections 3 and 4) was to give a definition that was applicable to the wide variety of formalized languages considered by mathematical logicians, and that gave natural results for each of those languages. The characteristic differences between so-called "logical languages," on the one hand, and languages with mathematical primitives and intended models with a domain of "nonlogical individuals," on the other, are explained and called upon at several points.

In Section 4 I consider the objection, raised by some commentators, that Tarski's definition does not require variation in the individual domain of interpretations when testing for logical consequence; if this is so, the definition would have several undesirable consequences. The objection is to some extent natural, since the definition, read word for word, says that a sentence X follows logically from a set of sentences K just in case all the interpretations of the extra-logical constants of K and X that make all sentences in K true also make X true (with no mention of changing the domain of individuals). However, I note, first, that Tarski clearly required domain variation in the test for logical consequence within languages containing extra-logical mathematical primitives; and second, I offer textual evidence to show that in these cases he assumed that the domain of the standard model or models was denoted by an extra-logical predicate, hence subject to reinterpretation in the test for logical consequence.

Section 5 recapitulates some of the motivations behind Tarski's proposal of his definition, with the purpose of showing their consistency and unity of purpose. To this effect, Tarski's motivations and their justification are set against the general logico-philosophical context in which the definition appeared. My final conclusion is that the definition successfully accommodated the main motivations that led Tarski to offer it.

2 Tarski's definition of logical consequence Tarski put forward his definition of logical consequence in a paper published in 1936 entitled "On the concept of logical consequence." The paper is a summary of an address given at the International Congress of Scientific Philosophy, held in Paris in 1935, a congress to which Tarski had been invited by Carnap (the paper appeared both in the proceedings of the congress and in a philosophical journal in Polish). It begins with some general remarks on the possibility of a precise definition of the concept of consequence. The essence of these remarks is that since the common concept is vague, it seems certainly difficult, and perhaps impossible, to reconcile all features of its use in the definition of a corresponding precise concept. Nevertheless, Tarski says, logicians had thought until recently that they had managed to define a precise concept that coincided exactly in extension with the intuitive concept of consequence. Tarski mentions the extraordinary development of mathematical logic in recent decades, which had shown "how to present mathematical disciplines in the shape of formalized deductive theories" ([30], p. 409). In these theories, consequences are extracted from axioms and theorems by rules of inference, "such as the rules of substitution and detachment" ([30], p. 410), of a purely syntactical (or "structural," in Tarski's word[1]) nature. "Whenever a sentence follows from others, it can be obtained from them—so it was thought—by means of the transformations prescribed by the rules" ([30], p. 410). According to Tarski, this belief of the logicians was justified by "the fact that they had actually succeeded in reproducing in the shape of formalized proofs all the exact reasonings which had ever been carried out in mathematics" ([30], p. 410).

But Tarski goes on to note that that belief of the logicians was wrong. There are some nonvague cases in which a certain sentence follows in the intuitive sense from a set of other sentences but cannot be derived from them using the accepted axioms and rules. These cases are provided by some ω-incomplete theories, theories in which for some predicate P all the sentences

A_0. 0 possesses the given property P,

A_1. 1 possesses the given property P,

and, in general, all sentences of the form A_n can be proved, but the universal sentence

A. Every natural number possesses the given property P,

cannot be proved on the basis of the accepted axioms and rules of inference. "Yet intuitively it seems certain that the universal sentence A follows in the usual sense from the totality of particular sentences $A_0, A_1, \ldots, A_n, \ldots$. Provided all these sentences are true, the sentence A must also be true" (Tarski [30], p. 411).

Tarski considers the possibility of adding an ω-rule to the accepted rules of inference, that is, a rule which allows us to deduce a universal sentence of the form of A from the set of sentences A_0, A_1, etc. However, he says that such a rule would differ in very essential respects from the old rules: it is not a finitary rule, whereas all the accepted rules in common deductive systems are finitary. Tarski then remarks that a certain finitary version of the rule could be stated for theories in which sufficient arithmetic can be constructed; let $B(P)$ be an arithmetical sentence of the theory coding (by means of a Gödel numbering) the statement that all the numeral instances of a predicate P can be proved in the theory by means of a certain antecedent set of rules; and let $C(P)$ be the universal generalization of that predicate (relativized to the natural numbers). Then the rule that $C(P)$ is deducible from $B(P)$ is a quite complicated to state, but finitary rule for the language of the theory in which arithmetic can be constructed, and we can add it to the antecedent set of rules of the theory. (Notice that the fact that all numeral instances of a predicate P are deducible by means of the antecedent rules does not imply by itself that a sentence $B(P)$ is deducible by the antecedent rules.) But Tarski immediately takes importance away from the suggestion of supplementing the old system of rules by means of a formalized finitary ω-rule. He points out that in view of Gödel's incompleteness theorem, no matter how many new finitary rules or axioms we add to the theory, and to successive finitarily supplemented extensions of the theory, it will still remain an incomplete theory, in fact an ω-incomplete theory. This discussion is enough to show that "in order to obtain the proper concept of consequence, which is close in essentials to the common concept, we must resort to quite different methods and apply quite different conceptual apparatus in defining it" (Tarski [30], p. 413).

The different methods and the different conceptual apparatus that Tarski has in mind are going to be "the methods which have been developed in recent years for the establishment of scientific semantics, and the concepts defined with their aid" ([30], p. 414; a footnote refers us to Tarski's monograph on truth). Tarski begins his preliminary analysis with two "considerations of an intuitive nature," which are: (a) when a sentence X follows from a class K of sentences, "it can never happen that both the class K consists only of true sentences and the sentence X is false" ([30], p. 414), and (b) this "consequence relation cannot be affected by replacing the designations of the objects referred to in these sentences by the designations of any other objects" ([30], p. 415). The justification for (b), according to Tarski, is that logical consequence is the same as *formal* consequence, "a relation which is to be uniquely determined by the form of the sentences between which it holds" ([30], p. 414). Tarski then says that (a) and (b) "may be jointly expressed" in a necessary condition for the relation of consequence to hold between K and X, which we quote in full:

> (F) If, in the sentences of the class K and in the sentence X, the constants—apart from purely logical constants—are replaced by any other constants (like signs being everywhere replaced by like signs), and if we denote the class of sentences thus obtained from K by 'K'', and the sentence obtained from X by 'X'', then the sentence X' must be true provided only that all sentences of the class K' are true ([30], p. 415).

(The 'F' of "condition (F)" most probably stands for 'form(ality)'.) After the statement of condition (F) Tarski adds, in parentheses:

> For the sake of simplifying the discussion certain incidental complications are disregarded, both here and in what follows. They are connected partly with the theory of logical types, and partly with the necessity of eliminating any defined signs which may possibly occur in the sentences concerned, i.e., of replacing them by primitive signs ([30], p. 415).

According to Tarski, if condition (F) were not only necessary for the relation of consequence to hold, but also sufficient, the problem of giving a satisfactory definition of consequence would have been solved, since "the only difficulty would be connected with the term 'true' which occurs in the condition (F). But this term can be exactly and adequately defined in semantics" ([30], p. 415). However, reflection shows that condition (F) is not in general a sufficient condition.

> This condition may in fact be satisfied only because the language we are dealing with does not possess a sufficient stock of extra-logical constants. The condition (F) could be regarded as sufficient for the sentence X to follow from the class K only if the designations of all possible objects occurred in the language in question. This assumption, however, is fictitious and can never be realized ([30], pp. 415-416).

(Tarski does not explain in the paper why the assumption "is fictitious and can never be realized." As we will see, other passages of the paper and of Tarski's contemporary work suggest that he has in mind certain denumerable languages in which arithmetic can be developed. None of these languages can contain "designations," primitive or defined, of all properties of natural numbers.)

It is also through semantics that, according to Tarski, we can overcome this difficulty. As he has said earlier, the main idea of the definition of logical consequence is not going to be original, and is implicit in the work of many contemporary and earlier logicians; but, as Tarski also has said, only using the recent precise definitions of the semantic concepts can that idea be developed in an exact manner (see [30], p. 414). To suggest how this can be done is the main purpose of Tarski's paper. The key idea is that instead of considering all replacements of extra-logical constants by other constants, we should consider all possible reinterpretations of the extra-logical constants. This can be done in a way satisfactory to Tarski with the help of the semantic notion of satisfaction, which he had just shown how to define in his monograph on truth. Using the notion of satisfaction, we can define the notion of model or realization of a class of sentences L in the following way. First we replace in a uniform way all the extra-logical constants in L by variables (of a corresponding grammatical category).

In this way we obtain a corresponding class of sentential functions L'. Then "an arbitrary sequence of objects which satisfies every sentential function of the class L' will be called a *model* or *realization of the class L of sentences* (in just this sense one usually speaks of models of an axiom system of a deductive theory)" ([30] p. 417). Using in turn the notion of model, Tarski defines the concept of logical consequence: "The sentence X *follows logically* from the sentences of the class K if and only if every model of the class K is also a model of the sentence X" ([30], p. 417).

Tarski immediately adds that the two characteristics of the notion of consequence incorporated in condition (F) can be shown to belong to the defined notion:

> …it can be proved, on the basis of this definition, that every consequence of true sentences must be true, and also that the consequence relation which holds between given sentences is completely independent of the sense of the extra-logical constants which occur in these sentences. In brief, it can be shown that the condition (F) formulated above is necessary if the sentence X is to follow from the sentences of the class K ([29], p. 417).[2]

Hence, as was to be desired, if the defined relation of logical consequence holds for a given pair $\langle K, X\rangle$, then also condition (F), the condition of formality, holds for it. (The proof is simple: suppose that X follows logically from K according to the definition; then there is no model of K which is not a model of X; so there is no substitution instance $\langle K', X'\rangle$ of $\langle K, X\rangle$ such that K' is true and X' is false; for if there were one such, it would readily provide a model of K that would not be a model of X.) Therefore, Tarski shows explicitly that his definition verifies at least this pretheoretic desideratum. (Although this can be shown, the converse cannot; that is, it cannot be shown that if X and K satisfy (F) then X follows from K according to Tarski's definition. But this is acceptable, since, as Tarski has already pointed out, (F) is not a sufficient condition for the ordinary notion of consequence.)

3 Theories of arithmetic and logical theories

3.1 The initial perplexity The example of ω-incomplete theories is the only example considered by Tarski in order to motivate his definition of logical consequence. According to Tarski, it seems intuitively certain that the universal sentence A follows "in the usual sense" from the set of sentences A_0, A_1, etc., and yet it is not derivable from them "by means of the normal rules of inference." This example is likely to cause some initial perplexity, for, under the usual first-order formalizations of arithmetic (where numerals are terms in which only extra-logical constants occur), a sentence of the form of A is not a model-theoretic consequence of a set of sentences of the forms A_0, A_1, etc. More perplexity can be added if we reflect on the fact that for first-order languages Gödel's completeness theorem assures us that the relations of model-theoretic consequence and of derivability by means of a few normal rules of inference coincide in extension. Etchemendy expresses this perplexity in passages like this: "…Tarski's examples involve the consequence relation for first-order languages, where the model-theoretically defined relation coincides with the syntactically defined relation. How can a semantic account be judged extensionally superior to the usual syntactic characterization if the two are, in fact, extensionally equivalent?" (Etchemendy [10], p. 85).[3]

Etchemendy's solution to this perplexity is "the flexibility Tarski allows in our choice of 'logical' constants. Clearly, if we choose to treat the numerals '0', '1', '2', ..., as logical constants, as well as the quantifier 'every natural number', then sentence A will come out a consequence of the infinite sentences $A_0, A_1, A_2, \ldots$; after all, any set that contains each natural number contains every natural number" (Etchemendy [8], p. 73; see also [10], p. 85). But immediately Etchemendy goes on to argue that Tarski's flexibility in the choice of logical constants results in some serious problems for his definition:

> Gödel sentences are a bit trickier,[4] due to their potential variety: all we can really say is that they will indeed come out as consequences of their corresponding theories if we treat all expressions of the language as logical constants. Unfortunately, this involves a certain trivialization of Tarski's analysis. For with this choice of logical constants, a true sentence is a logical consequence of *any* set of sentences whatsoever. This in fact points up a serious weakness in Tarski's account. It is clear that any given instance of the intuitive consequence relation can be made out to be a "Tarskian" consequence, at least on some selection of logical constants. But as soon as we extend this selection beyond the standard constants we also introduce many Tarskian consequences that are *not* instances of the intuitive relation ([8], p. 73).

Here Etchemendy adds a footnote: "This problem actually comes up even with the standard selection of logical constants. Note, for instance, that on Tarski's account ['$\exists x \exists y(\neg x = y)$'] is a logical consequence of any set of sentences, as long as '$\exists$', '$\neg$' and '$=$' are treated as logical constants" ([8], p. 73).

The correct solution to our perplexity is, however, more complicated. It does not consist in attributing to Tarski a nonstandard selection of logical constants in first-order arithmetic, but in observing that, when he gave his motivating example, he was not claiming that a version of the ω-rule in first-order arithmetic is valid. In the remainder of this section we will take a closer look at the sort of ω-incomplete theories that Tarski has in mind. A proper understanding of this matter will allow us to assess Etchemendy's claim that, if we are to make sense of Tarski's examples of intuitive cases of logical consequence, then his definition becomes trivial and even extensionally incorrect; to provide this assessment is one of the objectives of Section 4.

3.2 Logical languages and the formalization of arithmetic As a point of departure, it is necessary to call attention to the fact that Tarski intends his definition to be applicable to a wide variety of formal languages. As he says, he introduces his proposal as "a general method which, it seems to me, enables us to construct an adequate definition of the concept of consequence for a comprehensive class of formalized languages" ([30], p. 414). He of course thinks that his definition must be applicable to first-order languages, as well as to second- and higher-order languages of mathematical theories which contain primitives denoting mathematical notions, but he is also thinking of what we might call purely "logical" theories.

In a footnote, Tarski refers the reader "for a detailed description of a theory with this peculiarity [ω-incompleteness]" ([30], p. 410, note) to a paper of his entitled "Some observations on the concepts of ω-consistency and ω-completeness," published in 1933 (in the same journal where Gödel had published his papers on the completeness and the incompleteness theorems). We will not go over this paper in great detail, but we will describe briefly the theory for which the observations of the paper are made and to which Tarski refers in the paper on logical consequence. The theory is basically a simple theory of types. The symbols of its language are negation, the conditional, and the universal quantifier as logical primitives or constants, and infinitely many variables in each type. From these symbols formulas are constructed under certain appropriate type restrictions. In the theory there is a typical set of axioms for the propositional connectives and the quantifier, axioms of comprehension for each type, axioms of extensionality for each type, and an axiom of infinity. This axiom guarantees the existence of an infinity of objects of the lowest type, or individuals, but without specifying anything about their nature; this is spoken of in some works of the time as the "logical" sense of 'individual'. The rules of inference for the theory are substitution, detachment, universal instantiation, and universal generalization.

If we exclude the axiom of infinity, this theory is essentially the one that Gödel uses as a basis to formalize arithmetic in his incompleteness paper of 1931, and also essentially what Hilbert and Ackermann call the "calculus of levels" in the second edition of their logic manual; they use the theory also in a formalization of rational number theory in which real number theory can be developed (see Hilbert and Ackermann [16], pp. 121ff.; what they call the calculus of levels in the first edition of that book is what we would now see as a ramified theory of types with axioms of reducibility). But both Gödel and Hilbert and Ackermann use this simple theory of types as a basis for the formalization of these mathematical theories, not as a theory on its own. Thus, Gödel takes as the domain of individuals (the range of the variables of the lowest type) of his theory the natural numbers (hence, not individuals in a "logical" sense); Hilbert and Ackermann take the rational numbers. Gödel adds to the merely logical primitives of the calculus of levels some primitives of natural number arithmetic; Hilbert and Ackermann add some primitives of rational number arithmetic. And finally, Gödel takes the Peano axioms, including a second-order rendering of the axiom of induction, as primitive axioms of his theory; Hilbert and Ackermann take some suitable axioms for rational number arithmetic as primitive axioms of theirs.

In several works of this period Tarski too uses a simple type theory as what he calls a "logical basis" for the formalization of several different mathematical disciplines. In the resulting theories, mathematical primitives are superimposed on a more "logical looking" basis in which only the traditional logical primitives appear. However, in [24], to which he refers in the paper on logical consequence, Tarski takes the simple theory of types as a theory of its own, without arithmetical or any other mathematical primitives. In this theory, assuming that the axiom of infinity is of a logical nature, arithmetic can be developed within or reduced to "logic" in a well known fashion. For our purposes, it is enough to stress that in this theory only primitives of a distinctly logical nature appear (negation, the conditional, and the universal quantifier). Thus, the axioms, and in particular the axiom of infinity, are formulated with the help of only logical constants (and variables of all the types; in the case of the axiom of infinity only variables of the lowest types are needed). Arithmetical notions, including each numeral and the predicate 'to be a natural number', are defined with the help of only logical primitives, and so sentences which contain only arithmetical primitives come to contain, under the reduction, only logical primitives. Finally, the arithmetical axioms are derived as truths expressed in terms of the logically defined notions. A theory of this sort is ω-incomplete, and Gödel's incompleteness results apply to every consistent extension of it,[5] as Tarski points out at the end of [24]. Here he also advances in a sketchy form the idea that Gödel's investigations show that "the formalized concept of consequence will, in extension, never coincide with the ordinary one" ([24], p. 295).

In the footnote of [30] where he refers us to Tarski [24], Tarski also mentions a discussion of the topic in his 1933 paper on the concept of truth. The reference is to Section 5 of this paper, where Tarski presents what he calls the "general theory of classes" (a theory like the one of [24], i.e., a simple type theory with axioms of comprehension, extensionality, and infinity). It is precisely after noting that arithmetic can be developed within this theory (see [27], pp. 249ff.) that Tarski makes a shy remark in a footnote, submitting the idea that Gödel's results may show that the formalized concept of consequence does not coincide with the proper concept (see [27], p. 252).

The references to the theory of Tarski [24] and to the "general theory of classes" of [27] are not the only ones which reveal that Tarski, against what Etchemendy says, does not have in mind first-order formalizations of arithmetic in his paper on logical consequence. This point is confirmed by an accurate reading of other passages of the paper. Thus, in a footnote, Tarski notes that whereas already in [24] he had pointed out that the concept of consequence formalized by logicians did not coincide with the "proper concept" (precisely because of the phenomenon of ω-incompleteness), in that paper he "had expressed [himself] in a decidedly negative manner about the possibility of setting up an exact formal definition for the proper concept of consequence" ([30], p. 413, note 2). The reason for this was that, at that point, Tarski accepted only limited mathematical resources for use in metamathematical investigations. Namely, he accepted mathematical resources which enabled him to define the semantic concepts (truth, satisfaction, etc.) only for languages of finite order (first-order, second-order, etc.), but not for the language of the (simple) theory of types (which is, in fact, a language of infinite order; sometimes, as in the English translation of Hilbert and Ackermann [16] (Hilbert and Ackermann [17]), it is called the "calculus of order ω"). Tarski implies that now he accepts in metamathematics enough mathematics (in fact, enough set theory) to enable him to define truth, satisfaction, and therefore logical consequence for the full language of the theory of types. He explains in detail the changes in his point of view about this matter in the Postscript to the German translation of his monograph on truth (published in 1935); in this work, it is emphasized throughout that, in the metatheory adopted, the concept of truth can be defined only for languages of finite order, and the example for which the definition is given and most results are proved is a fragment of the simple theory of types in which the only quantifiers range over classes of individuals (cf. note 10 of the present paper).

Another indication that Tarski has in mind the "general theory of classes" when he writes the paper on logical consequence is provided by his parenthetical remarks after his statement of condition (F) (see Section 2 above); it seems that he intends the first remark to rule out the possibility that we replace a predicate by a predicate of a different type. The remark concerning the necessity of eliminating defined signs could be seen as a natural worry if Tarski were thinking of the arithmetical constants of his motivating example as being defined by means of logical constants; if they appear as abbreviations they should not be mistaken for mathematical primitives. No such worry would exist if Tarski were thinking of the numerals and the predicate ‘to be a natural number’ as primitives and logical constants.

It can be amply documented that in the works of this period Tarski reserves his most inclusive use of the word ‘logic’ for a system of logic based on the theory of types like the ones described above of [24] or [27]; such a “logic” is a system, therefore, in which arithmetical constants can be defined in terms of logical constants, and arithmetic developed as logic. Thus, in Tarski’s elementary logic manual, published in German in 1937 (revised, enlarged, and translated into English in 1941), we find remarks like the following:

> ‘…it turns out that the notion of number itself and likewise all other arithmetical concepts are definable within the field of logic. It is, indeed, easy to establish the meaning of symbols designating individual natural numbers, such as ‘0’, ‘1’, ‘2’ and so on. The number 1, for instance, can be defined as the number of elements of a class which consists of exactly one element. …Nor is it hard to define the general concept of a natural number: a natural number is the cardinal number of a finite class. We are, further, in a position to define all operations on natural numbers, and to extend the concept of number by the introduction of fractions, negative and irrational numbers, without, at any place, having to go beyond the limits of logic. Furthermore, it is possible to prove all the theorems of arithmetic on the basis of laws of logic alone (with the qualification that the system of logical laws must first be enriched by the inclusion of a statement which is intuitively less evident than the others, namely, the so-called *axiom of infinity*, which states that there are infinitely many different things). This entire construction is very abstract, it cannot easily be popularized and does not fit into the framework of an elementary presentation of arithmetic; in this book we also do not attempt to adapt ourselves to this conception and treat numbers as individuals and not as properties or classes of classes. But the mere fact that it has been possible to develop the whole of arithmetic, including the disciplines erected upon it—algebra, analysis and so on—, as a part of pure logic, constitutes one of the grandest achievements of recent logical investigations” ([32], p. 81; the English text is an almost literal translation of the German text, for which see [31], pp. 50–51).

In this passage it is very clear what is Tarski’s most comprehensive use of the term ‘logic’, at least in this period. His emphasis on the definability in logical terms of the numerals and of the concept of natural number is important for us to note. It is also interesting to note Tarski’s remark that in his book he does not present mathematical theories in the logical fashion for reasons of simplicity and pedagogy.

3.3 Logical constants and arithmetical constants Also in Tarski’s manual, we find very significant remarks about the distinction between logical and arithmetical constants (the distinction, that is, when the arithmetical constants are taken as primitives, for reasons of simplicity or other reasons, as in Tarski’s elementary manual or in Gödel [12]). In a section entitled “The specifically mathematical expressions and the logical expressions; mathematical logic” in the German edition, and “Logical constants; the old and the new logic” in the English edition, Tarski says:

> The constants with which we have to deal in each mathematical discipline may be divided in two big groups. The first group consists of the expressions of a

> specifically mathematical character. In the case of arithmetic, for example, they are terms denoting either individual numbers or whole classes of numbers, relations between numbers, operations on numbers, etc.…In mathematical statements, however, there are also expressions of a much more general character, expressions which are met constantly both in considerations of everyday life and in every possible field of science, and which represent an indispensable means for conveying human thoughts and for carrying out inferences in any field whatsoever; such expressions as 'not', 'and', 'or', 'is', 'every', 'some' and many others belong here. There is a special discipline, namely *logic*, considered the basis for all other sciences, whose concern it is to make precise the content of such concepts and to lay down the most general laws in which these concepts are involved ([31], p. 12, my translation; for the text of the English edition, which I have followed as closely as possible, see [32], p. 18).

This text adds support to the view that Tarski would not be naturally thinking of including arithmetical expressions as logical constants when they appear as primitives of a formalization of arithmetic. This is, in fact, the standard terminological usage. The same distinction appears implicitly in a note in Tarski's paper on the concept of truth. Here, after presenting the "general theory of classes" (see above), he compares it with Gödel's system "P" of [12] (see above), and says: "Apart from certain differences of a 'calligraphical' nature, the only distinction [between the system P and the general theory of classes] lies in the fact that in the system P, in addition to three logical constants, certain constants belonging to the arithmetic of the natural numbers also occur" (Tarski [27], p. 248).

In [30], Tarski leaves as an open problem the question of characterizing the sets of logical constants of those languages to which his definition of logical consequence is applicable (see [30], pp. 418–419). The problem is a problem because the applicability of Tarski's definition to a particular language presupposes a previous division of the terms of the language into logical and extra-logical. However, this situation is tolerable, and the need for such a definition is not urgent, because, as Tarski says in a letter of 1944, "it is clear that for all languages which are familiar to us such definitions [of 'logical term' and 'logical truth'] can be given (or rather: have been given); moreover, they prove fruitful, and this is really the most important. We can define 'logical terms', e.g., by enumeration" ([33], p. 29).

In 1966, in a lecture that would not be published until after his death, Tarski returned to this problem, advancing a tentative definition of the concept of logical constant. I will not discuss Tarski's proposal of 1966 in detail here, but I will make some remarks relevant to our current discussion. As one might expect, this definition has three important properties: (1) it is generally applicable to the wide range of languages to which Tarski's definition of logical consequence is applicable, (2) it generates the usual extension of logical constants for a language like that of *Principia*; and (3) it generates the usual extension of logical constants in languages with mathematical primitives and intended models whose universes of discourse contain "nonlogical" individuals. Tarski suggests that we call logical notions of an interpreted language those notions invariant under all one-one transformations of the universe of individuals onto itself (see [34], p. 149). The first language used by Tarski to test the appropriateness of the definition is, not surprisingly, the language of the theory of types of *Principia Mathematica* (see [34], pp. 150–151). In [26], originally published in 1935, Tarski and Lindenbaum had proved that all the notions which can be defined in *Principia* are invariant under all permutations of the universe of individuals.[6] That is, the classes of individuals, relations of individuals, classes of classes of individuals, etc., which can be defined in the theory of types stay the same after any permutation of the universe of individuals. In general, Tarski and Lindenbaum's theorem guarantees that all mathematical notions definable in the logicist fashion in the theory of types are logical notions.

However, when the proposed definition is applied to formalizations of mathematical theories with undefined mathematical primitives and intended models whose universes of discourse contain "non-logical" individuals, it declares extra-logical the notions denoted by these primitives. Tarski's example is set theory formalized in first-order with a single primitive predicate for membership and a universe of discourse to which all sets belong (together with other individuals, or *Urelemente*; see [34], p. 153). Obviously membership is not invariant under all permutations of a universe of individuals which contains all sets, so it is not declared a logical notion by Tarski's proposed definition. (A unary primitive predicate whose intended extension was the class of all sets would not be logical either, provided that the domain of individuals contained at least some thing which is not a set.) If we apply Tarski's definition to a typical first-order formalization of arithmetic, we obtain the result that individual notions, like the number 0, and unary functions other than the identity function, like the successor function, are not declared logical. Hence none of the numerals in a standard first-order formalization of arithmetic are logical constants under Tarski's proposal. (A unary primitive predicate whose intended extension is the set of natural numbers would not be logical either, provided that the domain of individuals contained some objects other than natural numbers.)

3.4 Conclusion If the arithmetical expressions are not logical constants when they are primitives of our formalization of arithmetic, A will not be declared a logical consequence of the set of sentences A_0, A_1, etc. by Tarski's definition (many interpretations of 'to be a natural number' and of '0', '1', '2', etc. will make all the sentences of the infinite set true but A false). The solution suggested by the textual evidence is that when he gives his motivating example Tarski is not thinking of the arithmetical expressions as primitives, but as defined terms; defined, that is, with the help of logical constants, within the framework of a sufficiently powerful logical theory. If the predicate 'to be a natural number' and the numerals '0', '1', '2', etc. are defined in the logicist fashion within the framework of an appropriate logical theory, A will follow from the set of sentences A_0, A_1, etc. according to Tarski's definition, since the only extra-logical constants subject to reinterpretation in the test for logical consequence will appear, if there are any, in the predicate P. And any class that contains all the logically defined "numbers" will include the class denoted by the logically defined predicate 'to be a natural number', since this will be no other than the class of all the logically defined "numbers."[7]

Let us conclude this section by considering a seeming objection that might be posed to Tarski's informal conception of the logical correctness of the ω-rule. This objection might be put as follows: if the inferences licensed by a version of the ω-rule are correct, then they should be declared logically correct under any formalization of arithmetic, whether it is a formalization in the logicist fashion or a first-order formalization with the numerals and the predicate 'to be a natural number' as extra-logical primitives; hence, the objection would conclude, either Tarski's definition is wrong or it must resort to nonstandardness in the choice of logical constants, since it does not declare logically correct those inferences in a first-order formalization.

This objection is no better grounded than the false thesis that a correct argument should be logically correct under any of its formalizations. It is a familiar point from introductory logic courses that some correct arguments are logically correct under some formal renderings but not under some other, less discriminatory ones. 'All Greeks are human and all humans are mortal, so all Greeks are mortal' is a correct argument, but only some of its formal renderings are logically correct; e.g., the argument with premise 'p & q' and conclusion 'r' in a propositional language is not logically correct, but the argument with premise '$\forall x(Gx \to Hx)$ & $\forall x(Hx \to Mx)$' and conclusion '$\forall x(Gx \to Mx)$' in a first-order language is logically correct. This is a good way to see that neither we nor Tarski are under any constraint to accept the thesis that if one version of the ω-rule is logically valid then all its versions, and in particular first-order versions, are logically valid as well. And there is no indication at all that Tarski ever considered such a thesis. We have seen that, as a matter of fact, the logical validity of the ω-rule is asserted by Tarski only when he discusses versions of it in the simple theory of types. Tarski did not think that first-order versions of the ω-rule were logically valid. He thought only that some of its versions in the simple theory of types were.

4 Truth in all domains and the domain of logic

4.1 The problem of domain variation We have seen that one of the perplexities likely to appear when reading [30] is created by Tarski's insistence on the idea that the proper concept of consequence cannot coincide with a syntactically defined concept of consequence, whereas Tarski should have known perfectly well that Gödel's completeness theorem guarantees the coincidence of the model-theoretic and the syntactic concepts for first-order languages. We have also seen how this perplexity disappears if we note that Tarski is not trying merely to define the notion of consequence for sentences of first-order languages, but also for other languages, and in particular, in his motivating example, for a quite powerful logical theory in which a great deal of mathematics can be developed; here, it is also a result of Gödel, namely, the first incompleteness theorem, that guarantees that the syntactically defined concept cannot exhaust the model-theoretic concept.

Therefore Tarski does not reserve his most inclusive use of the term 'logic' simply for higher-order logic, as we understand this concept today, but for a theory with more powerful assumptions, especially an axiom guaranteeing the infinity of the domain of "logical" individuals (of which he speaks as less evident than the others). As we have seen and we will still have occasion to see, the acceptance of this strong "existence assumption" as a truth of logic in a certain use of the term 'logic' does not seem to trouble Tarski excessively. But all this leaves us wondering what Tarski's position was at the time regarding the properties of logical consequence and logical truth for mathematical theories couched in first-order, second-order, and in general finite order languages, in which the notion of individual is not logical and undefined mathematical primitives appear; for the standard model-theoretic definitions of these properties are such that no sentence is declared a logical truth which asserts the existence of a number of individuals greater than one.

Etchemendy repeatedly insists on a point that we should consider here. Tarski's definition of logical consequence in [30] declares logically true (or, logical consequences of any set of sentences) those sentences which are true for all interpretations of their extra-logical constants. No mention is made of changing the domain of individuals when testing for logical truth or logical consequence.[8] Leaving the domain invariant in the test would have the effect that many assertions about the cardinality of the domain of the intended interpretation of the language, which can be expressed with the help of just logical constants, would be declared logically true by Tarski's definition (read word for word). We saw already that Etchemendy gives the example of the first-order sentence '$\exists x \exists y(\neg x = y)$', which, if Etchemendy's point is right, will be declared logically true as long as our point of departure in the test for its logical truth is a language whose intended interpretation includes a domain of two or more individuals.

This, according to Etchemendy, is a "flaw" of Tarski's definition of which Tarski was "surely aware" (see Etchemendy [8], p. 73). Also according to Etchemendy, this supposed flaw was "no doubt partly responsible" for Tarski later "giving up" his definition of logical consequence of 1936. Etchemendy infers that Tarski gave up this definition from the fact that in later work (viz. [35]) he gives the standard model-theoretic definition of logical consequence for first-order languages, including an explicit mention of changing the domain of the interpretations. But it is extremely doubtful that Tarski saw himself as proposing different definitions in [30] and [35], for here he refers the reader again to [30] and [27] "for formal definitions and a detailed discussion of semantical notions (satisfaction, truth, logical consequence, logical truth)" ([35], p. 8, note 7). Despite this, Etchemendy pictures Tarski as choosing to define logical consequence in 1936 without requiring domain variation. And this even in view of the fact that, as Etchemendy himself says, Tarski could not have failed to notice that in this way his definition would differ essentially from that "presupposed" by Gödel's proof of the completeness theorem for first-order logic (and also by Hilbert and Ackermann, who described the problem of finding such a proof and proposed it as an open problem in Hilbert and Ackermann [15] (see p. 68)).

4.2 Truth in all domains and logical truth A more balanced appreciation of the real situation concerning this set of issues is again more complicated than the one Etchemendy's reading provides. We may start by repeating the simple point that first-order logic is only a very limited part of what Tarski is thinking of as logic when he writes the paper on logical consequence. The same is true for Hilbert and Ackermann, who are among the first to single out first-order logic, but only as a simple and useful *fragment* of logic; logic is also taken by them to include the calculus of levels.[9] Probably the same could be said of Gödel; in any case, his proof for first-order languages that the notion of a sentence true in every individual domain coincides extensionally with that of a sentence derivable in a certain calculus, cannot have been seen at the time as a proof that the notion of provable sentence coincides with that of logically true sentence; for this latter concept was taken to be more inclusive than the concept of sentence of first-order logic true in every individual domain. Furthermore, although many of the sentences that Tarski and Hilbert and Ackermann are willing to accept as logically true in higher-order languages are true in every individual domain, they are ready (with caution) to accept as a logical truth within the framework of the "general theory of classes" a higher-order sentence containing only logical constants which effectively states that the cardinality of the domain of "logical" individuals is infinite (as we saw and we will again see in Tarski, and can be seen in Hilbert and Ackermann [15], p. 88). From this axiom of infinity, any first-order sentence asserting that the domain of logical individuals contains at least a certain finite number of elements would follow logically, and therefore would be generously considered logically true by Tarski within the context of a logical theory like the general theory of classes.

However, it is extremely important to note that Tarski was also perfectly aware of the idea that the term 'logic' might be understood so that there would not be among logical truths any assertions about the cardinality of the universe of individuals (even of "logical" individuals). We find some illuminating remarks of Tarski about this point in his article on truth, at the end of Section 4, in which he has discussed in a general setting "the concept of true sentence in languages of finite order." Here he speaks sympathetically of understanding logic in this more restricted fashion, in the context of a discussion of deductive theories which are "parts of logic," or fragments of logic (see [27], p. 239). He first points out that

> [in sciences that are part of logic] the concept of correct sentence in every individual domain…deserves special consideration. In its extension it stands midway between the concept of provable sentence and that of true sentence; the class of correct sentences in every domain contains all theorems and consists exclusively of true sentences.… This class is therefore in general narrower than the class of all true sentences; it contains, for example, no sentences whose validity depends on the magnitude of the number of all individuals.… If it is desired to transform the system of the provable sentences of every science [which is a part of logic] into a complete one, it is necessary at the outset to add sentences to the system which decide the question how many individuals exist ([27], p. 240).

Tarski is observing here that when we take a theory which is a part of logic, the class of its sentences true in every individual domain will in general (but not always, as Tarski will soon point out) be more inclusive than the class of provable sentences, and less inclusive than the class of true sentences. Although it is far in some respects from the example Tarski has in mind, we can give as an example of one such theory the pure second-order language of identity without extra-logical primitives;[10] it is relevant to remark that this theory, like the one Tarski has in mind, would be taken by Tarski to be a purely logical theory, a part of logic. In this language many sentences which are true in every individual domain are not provable on the basis of a *prima facie* fairly comprehensive set of axioms and rules of inference for second-order logic (for example, it is well known that there is a sentence of this language which is true in every individual domain if and only if the continuum hypothesis is true, and another which is true in every individual domain if and only if the continuum hypothesis is false; neither sentence is provable in a standard calculus for second-order logic, but one of them must be true in every individual domain). A sentence of this language like '$\exists x \exists y(\neg x = y)$' is a true sentence about "logical" individuals that is not true in every individual domain. In the last sentence of the text just quoted, Tarski is saying that in order to make complete a theory (or "science") which is part of logic, it is necessary to add to it sentences of its language which decide the cardinality of the domain of individuals; otherwise, sentences that can be formulated only with the help of logical constants, like '$\exists x \exists y(\neg x = y)$', may be undecidable in the theory investigated (neither it nor its negation will be provable). And he continues:

> But for various reasons another point of view seems to be better established, namely the view that the decision regarding such problems should be left to the specific deductive sciences, whilst in logic and its parts we should try to ensure only that the extension of the concept of provable sentence coincides with that of correct sentence in every individual domain. For a supporter of this standpoint the question whether the extension of these two concepts is actually identical is of great importance. In the case of a negative answer the problem arises of completing the axiom system of the science studied in such a way that the class of provable sentences thus extended now coincides with the class of sentences which are correct in every domain. This problem, which properly is equivalent to the question of structurally characterizing the latter concept, can be positively decided only in a few cases…[footnote:] In the case of the lower functional calculus this problem, which is raised in Hilbert and Ackermann [15], p. 68, has recently been decided by Gödel, see [11] ([27], p. 240).

Tarski is here speaking pretty sympathetically of the view that logic ("and its parts") should care only to prove sentences which are true in every individual domain (but not endorsing the view completely, as we are about to see; notice also the "seems" in the first sentence, and the noncommittal reference to "a supporter of this standpoint"). According to this view, sentences with implications about the cardinality of the domain of individuals (even if they contain only logical vocabulary) are to be postulated or proved *only* within the special sciences formalized with the help of logic. Tarski also notes that in only a few cases the problem of the coincidence of the notion of sentence true in every individual domain and that of provable sentence can be solved in the affirmative. We know, for example, that the answer is negative for the pure second-order language of identity. And significantly, Tarski notes that one *case* in which the answer is affirmative is the calculus of first-order logic, inserting a reference to Gödel's 1930 article.

Despite these sympathetic remarks, we have only to turn one page in Tarski's paper on the concept of truth to find him introducing the same system of logic of [24], already described above; that is, a simple theory of types with axioms of comprehension, extensionality, and infinity. Here this theory is called, as we have seen, "general theory of classes," and is characterized as "noteworthy because, in spite of its elementary structure and its poverty in grammatical forms, it suffices for the formulation of every idea which can be expressed in the whole language of mathematical logic" ([27], pp. 241–242). That is, the general theory of classes, or the theory of [24], comprehends all the parts of logic; it can, in fact, be identified with logic. About the axiom of infinity, we find the following remark, in a footnote: "In adopting the axiom of infinity we admittedly give up the postulate according to which only the sentences which are correct in every individual domain are to be provable sentences of logic [here Tarski asks the reader to see the text quoted above (i.e., [27], p. 240)]" ([27],

p. 243).[11]

On the logical status of the axiom of infinity there were many discussions at the time. Probably the most extended view was that it was not a truth of logic, properly speaking; this was Russell's view, which had led him to state in conditional form the theorems of *Principia* that depend on a postulate of infinity, with such a postulate as antecedent. But as we have seen, many mathematical logicians of the time did not have much trouble accepting the postulate properly as an axiom of logic. Moreover, philosophers of a logicist persuasion exhibited a natural self-serving tendency to license its inclusion among logical principles. Thus Hempel, in his 1945 paper: "the axiom of infinity does not belong to the generally recognized laws of logic; but it is capable of expression in purely logical terms and may be treated as an additional postulate of logic" (Hempel [14], p. 389); or Carnap, who in [1] puts forward an interpretation of his "language II" according to which existential assumptions like the axiom of infinity assert not the existence of a certain quantity of objects, but of "positions," the existence of a certain quantity of positions being a logical question (see Carnap [2], pp. 140ff.); or Ramsey, who in his 1925 paper interpreted Wittgenstein as having shown that assertions about the cardinality of the universe of individuals are either tautological or contradictory (although for most of them, including the axiom of infinity, we do not know what their status is), which gives us all reason to believe that the axiom of infinity is a tautology (see Ramsey [21], pp. 59ff.).

All these considerations suggest that Tarski's viewpoint was more sophisticated than Etchemendy concedes. In particular, it does not seem charitable enough to picture Tarski as believing that sentences expressing facts about the cardinality of the domain of individuals are not logical truths, realizing that his definition of logical consequence declares some such sentences logical truths (hence that it was "flawed"), and despite all this still going ahead and proposing the definition. It is more accurate to see Tarski as proposing a definition sufficiently general to be able to include as a special case a notion of logical consequence and logical truth established by the normal usage of mathematical logicians. This was a usage with which he had no excessive problems in agreeing, according to which the theorems of a powerful logical theory like the general theory of classes come to be considered truths of logic. However, as we have seen, Tarski was also perfectly aware that logic, and its fragments, could be understood as theories without cardinality assumptions. The theory of classes, in particular, could be stripped of the axiom of infinity. The resulting theory would be such that all its theorems would be true in all individual domains.

4.3 Mathematical theories and domain variation Tarski naturally intended his definition to be applicable not only to purely logical theories, but also to mathematical theories with special mathematical primitives and postulates. In the works of this period Tarski considers several mathematical theories formalized using a logical apparatus, or 'logical basis', to use Tarskian terminology, without any cardinality assumptions. Generally, this logical basis for formalization is again the calculus of levels, but without the axiom of infinity (this is also, as we saw, the "logic" of Gödel [12]); sometimes it is the calculus of first-order. In these formalizations: (1) the domain of objects of the mathematical theory is either identified with the domain of individuals or taken to be defined by a primitive extra-logical predicate (which does not apply to other members of the individual domain); and (2) besides purely logical constants, there are some mathematical primitives denoting objects in the intended domain of mathematical objects or relations on the intended domain of mathematical objects (for examples in which the logical basis is the calculus of levels see Tarski [23], originally published in 1931, and [26]; for examples of theories formalized in first-order, see Tarski [25] and [29], the latter originally published in 1936). These languages are not purely logical, for they contain undefined extra-logical primitives; and they do not assume a purely "logical" notion of individual, since the objects of the intended domain of mathematical objects are taken to be objects of a certain nature. In these cases Tarski does not take any assumptions about the cardinality of the domain of individuals to form part of the underlying logic of the theory alone, although of course the properly mathematical axioms of the theory may constrain the cardinality of its possible interpretations. This is only natural, because it is intuitively not the business of the underlying logic to decide how many individuals of a certain nature (natural numbers, real numbers, points) there are.

If we follow Etchemendy, we are led to conclude that Tarski's definition declares logically true sentences like '$\exists x \exists y(\neg x = y)$' or a formulation of the axiom of infinity in the theory of types, which contain only logical constants, provided that the *intended interpretation* of the language involves a domain with a sufficient number of individuals. This is so, according to Etchemendy, because Tarski's definition does not contemplate reinterpretations of the language in which the domain of individuals varies *with respect to the domain of the intended interpretation*. Etchemendy criticizes this consequence of the definition on the basis of a tacitly subscribed view of logic, according to which assertions expressing facts about domain cardinality cannot be logical truths. But if Etchemendy's were a correct reading, the definition would have an even more serious defect. The definition would lead to contradiction. Suppose that one day we are engaged in studying, with our logic without prior cardinality assumptions, a certain interpretation of a set of axioms that admits of intended interpretations of different cardinalities, and also that this fact can be "expressed" in the language of the theory. To be more definite, suppose that we have a first-order formalization of the axioms for Boolean algebra, and that we are studying a Boolean algebra of two elements which we take to be an intended interpretation for the axioms. Applying Tarski's definition we would conclude that the sentence '$\exists x \exists y \exists z(\neg x = y \,\&\, \neg x = z \,\&\, \neg y = z)$' *does not follow* from the set of axioms for Boolean algebra. Suppose that the day after we are studying an infinite Boolean algebra. Then applying Tarski's definition we would conclude that '$\exists x \exists y \exists z(\neg x = y \,\&\, \neg x = z \,\&\, \neg y = z)$' *follows* from the set of axioms for Boolean algebra. Tarski's definition would not be merely flawed; it would be useless, under any conception of logical truth.

The crucial point that we must acknowledge here is, of course, that sentences like '$\exists x \exists y(\neg x = y)$' and '$\exists x \exists y \exists z(\neg x = y \,\&\, \neg x = z \,\&\, \neg y = z)$', or an axiom of infinity expressed with purely logical constants, mean different things in interpretations with different individual domains. In one interpretation '$\exists x \exists y(\neg x = y)$' will mean that there are two natural numbers, in another that there are two real numbers, and so on. Only if the intended interpretation of the theory has as its domain the collection of all "logical" individuals (of the world) will '$\exists x \exists y(\neg x = y)$' mean "there are two things," and in this case we have seen that Tarski and others would be happy to call it a logical truth. But clearly sentences meaning "there are two natural numbers," "there are two real numbers," and so on, should not be logical truths of the underlying logic for axiomatic first-order theories of the natural numbers, the real numbers, and so on. Etchemendy's argument, however, can be used to establish that Tarski's definition declares the corresponding sentences logically true, since Tarski's definition, Etchemendy argues, does not allow for variation of the individual domain in the test for logical consequence.

But this objection to Tarski's definition would be out of place if Tarski had in mind mathematical theories in whose canonical formulation the domain of objects of the intended interpretation or interpretations is the extension of a primitive predicate of the language of the theory. Assuming that to specify a domain for an interpretation is nothing but to give an interpretation for such a predicate, Tarski's definition would allow for domain variation in the test for logical consequence. Under this assumption, sentences like '$\exists x \exists y(\neg x = y)$' would not be declared logical truths, because they would be mere unofficial abbreviations for other sentences with relativized quantifiers ('$\exists x \exists y(Nx \;\&\; Ny \;\&\; \neg x = y)$', for example). It is natural to picture Tarski as having in mind the idea that the domain of the intended model or models is denoted by an extra-logical predicate, but without even thinking of formulating or caring to formulate this as an explicit requirement for the applicability of his definition.

A look at Tarski's logic manual [31], where the importance of making clear all unofficial assumptions (to the beginner) is especially obvious, shows that indeed Tarski had in mind such a canonical formulation of mathematical theories and, furthermore, that he saw domain variation in the test for logical consequence as change in interpretation of a primitive predicate. Thus, he sets up a very simple first-order theory for pedagogical purposes of illustration (Tarski calls it a "miniature theory"), whose language has the symbols '**S**' and '$\cong$' as primitives. "The former is an abbreviation of the term 'the set of all segments'; the latter designates the relation of congruence" ([32], pp. 120–121; see also [31], pp. 84–85). The axioms of the theory are: "Axiom 1. For any element x of the set **S**, $x \cong x$.... Axiom 2. For any elements x, y and z of the set **S**, if $x \cong z$ and $y \cong z$, then $x \cong y$" ([32], p. 121; see also [31], p. 85). Tarski then describes, with words reminiscent of [30], the first step on the way to the notion of a model of such a theory: turning its axioms and theorems into sentential functions.

> Let us replace the primitive terms in all axioms and theorems of our theory by suitable variables, for instance, the symbol '**S**' by the variable 'K' denoting classes, and the symbol '$\cong$' by the variable 'R' denoting relations.... The statements of our theory will then be no longer sentences, but will become sentential functions which contain two free variables, 'K' and 'R', and which express, in general, the fact that the relation R has this or that property in the class K ([32], p. 122).

Tarski goes on to introduce the notion of a model of a theory, not as a general notion, but using his "miniature" theory as an example:

> If a relation R is reflexive and has the property **P** in a class K [i.e., that for any elements x, y and z of the class K, if xRz and yRz, then xRy], we say that K and R together form a *model* or a *realization of the axiom system* of our theory, or, simply, that they satisfy the axioms ([32], p. 123; see p. 122 for "property **P**").

In the next pages, Tarski shows that Axioms 1 and 2 have different models and how using some of these one can show that a certain sentence does not follow from the axioms.

We should remember at this point that just after defining the notion of a model of a set of sentences in his paper on logical consequence, Tarski says, using almost the same words he uses in his manual, that "in just this sense one usually speaks of models of an axiom system of a deductive theory" ([30], p. 417). This "usual sense" of 'model', which can be abundantly illustrated with many examples taken from Tarski's contemporary work, is the usual sense of 'model' according to which one specifies a model for a theory when one specifies both a *domain* and certain objects and relations *in the domain* as the meanings of the primitive symbols of the theory, and it is also, as we have just seen, the notion of model introduced in Tarski's manual.

It can be amply documented that many of the theories containing mathematical primitives that Tarski formalizes (with the help of a logical basis without cardinality assumptions) explicitly contain extra-logical predicates which are true of all the objects in the domain of the intended interpretation (or interpretations, in algebraic theories) of the theory. All the first-order theories Tarski gives in the examples of [31] contain a primitive symbol for the universe of discourse (see [31], pp. 84ff., and [32], pp. 120ff.), whose interpretation varies each time that a new model of the theory is given. The (second-order) theory of real arithmetic he describes in [31] contains the primitive predicate 'Zl' (for the German '*Zahl*'), meaning intuitively "to be a real number" (see [31], pp. 98ff.). In [28], a 1935 paper on Boolean algebra, the language of the theory contains a primitive predicate 'B' which Tarski reads as "the universe of discourse." Similar predicates appear in other papers on algebraic theories, where the multiplicity of intended models of the theory (with different domains of different cardinalities) is the natural theoretical situation.

We also must not forget other examples of contemporary investigations conducted by Tarski in which results about the class of all interpretations of a theory are proved. Of special significance is, I believe, the brief note of 1934 in which an anonymous editor of *Fundamenta Mathematicæ* announces on behalf of Tarski, without proof, what is now known as the upward Löwenheim-Skolem-Tarski theorem for first-order languages, and other results obtained by Tarski and presented in a seminar taught by him in 1927–1929.[12] In this note the author mentions two examples of theories formulated in first-order languages whose intended interpretations have denumerable domains, but that also possess other nonisomorphic, denumerable models in which the domain is different from that of the intended interpretation.[13] The third result stated in the note is the "upward" theorem, which guarantees the existence of models of all infinite cardinalities for first-order theories with countably infinite models; therefore, and hence its special significance for us, a theorem asserting the existence of models of different cardinalities for one and the same first-order mathematical theory. If Tarski is working with the "usual sense" of model when he proves the results of [25] and when he writes [30], it is not natural to think that in one case he allows models of different cardinalities for a theory and in the other case he does not.[14]

These are very strong considerations in favor of a reading "between the lines" of the definitions of model and logical consequence in Tarski [30]. It should not be forgotten that the paper is a summary of an address to a philosophical audience, and that

Tarski may not have made the effort to be precise in a point that he probably took for granted.[15] The supposition that Tarski did not contemplate domain variation is nearly impossible to reconcile with his contemporary work. A more explanatory supposition is that he *did* contemplate domain variation, but was not sufficiently explicit about it, for reasons which can be understood in view of the preceding explanations.

5 Tarski's definition in context The considerations of the preceding two sections help us understand some of the main motivations behind Tarski's proposal of his definition of logical consequence. One of these motivations was the need to offer a definition generally applicable to the formalized languages familiar to mathematical logicians, that agreed as much as possible with the "ordinary usage" among them of the notion of logical consequence. This involved giving a definition that could be applied to mathematical theories formalized in first-order languages, in second- and higher-order languages, and in a language with infinitely many types of variables, like the language of the general theory of classes. His definition accomplished this, since the notions involved in its application to these languages (satisfaction, model, logical constant) could be precisely defined (or, in the case of the notion of a logical constant, characterized by enumeration of the logical constants, following established usage). And it accommodated fairly well ordinary mathematical use of informal notions; in particular, it did not require any nonstandard treatment of domain variation in its application to, e.g., mathematical theories formalized in first-order languages with mathematical primitives.

Another important motivation for Tarski derived from the fact that some of the languages to which the definition should be applicable, like the language of the general theory of classes, were thought to be sufficiently powerful to codify and develop all of classical mathematics. A test for the adequacy of a definition of logical consequence would be whether, for the language of the general theory of classes, it declared consequences of all sets of sentences all the true sentences of classical mathematics (or, to be more exact, the sentences which, under the appropriate "translation," express truths of classical mathematics). This test could not be passed by any "syntactic" definition of logical consequence, in view of Gödel's incompleteness theorem, applicable to the general theory of classes. That is why Tarski stresses that for all appropriate languages, the notion of a sentence true in all interpretations (or, equivalently, of a sentence that follows from every set of sentences) coincides in extension with the notion of an analytical sentence, as this term had been defined in "general syntax" by Carnap (see [1], pp. 135ff., or [2], pp. 182ff.; for Tarski's statements about the equivalence of truth in all interpretations and "analyticity," see [30], p. 418). Carnap's definition guaranteed that for all appropriate languages, like his "language II," all the truths of classical mathematics are analytical, and the proof of the equivalence of the notion of analyticity thus precisely defined with the notion of truth in every interpretation must probably have been seen by Tarski as bearing philosophical significance.

That this was one of Tarski's motivations may come as a surprise, since the general view nowadays is that any theory powerful enough to develop some substantive mathematics (not to say all of classical mathematics) does not deserve the title "logic," and its truths do not deserve the title of "logical truths" (or of "logical consequences" of any set of sentences). In some cases, this general view may even have been fostered by Gödel's incompleteness theorem, since it is perhaps a reasonable stipulation on the meaning of the expression 'logical truth' that the set of logical truths of a language like the language of the general theory of classes must be recursively enumerable; and it is one of the consequences of Gödel's theorems that the set of sentences of this language true in an interpretation with an infinite basis of individuals is not recursively enumerable.

But we know that Carnap reacted to Gödel's discoveries by concluding not that they meant an objection to the logicist program broadly conceived, but that they made untenable the identification of the concept of logical truth with an effective concept of deducibility or derivability within some fixed formal system. That is why, in part under the influence of Tarski, Carnap developed in his [1] definitions of noneffective concepts which, in the appropriate languages, included in their extension all the (translations of the) truths of classical mathematics. Instead of thinking that Gödel's results showed that not all mathematical truths are (under some appropriate translation) logical truths, Carnap and Tarski thought that they showed that the notion of logical truth is not an effective notion.[16] The opening pages of Tarski [30] should be read in this light.

Carnap's definition of analyticity in [1] was not intended to capture the general notion of "truth by virtue of meaning" in natural language; this concern will appear in Carnap only in subsequent years. It was not a concern of Tarski either. Etchemendy's misinterpretation of the example of sentence A as a consequence of the set of sentences A_0, A_1, etc. is in part dictated by the assumption that Tarski's definition intends to capture the notion of "implication by virtue of meaning" of the expressions selected as logical constants (in the example, the numerals and the predicate 'to be a natural number'). Our interpretation showed that Tarski did not think of this as an example of consequence in virtue of the meaning of the expressions involved (an example of analytical entailment, among many others of ordinary language), but as an example of logical consequence in a certain strict sense (namely, in the sense of consequence in which only the logical constants of *Principia Mathematica* appear essentially). It is important to stress the fact that Tarski was not trying to capture with his definition the notion of implication by virtue of meaning, since it might perhaps be easy to criticize the adequacy of Tarski's definition relying on an identification of the "proper concept" of logical consequence with the concept of analytical implication.

However, Carnap's justification for using the loaded term 'analytical' for his defined notion in [1] was obviously that all the sentences falling under its extension were supposed to possess certain epistemic or modal properties (to be *a priori*, to say nothing about the world, etc.). It is unlikely that Tarski ever thought that a justification for his definition might lie in the fact that it delimited some epistemically or modally privileged set of sentences.[17] We have the testimony of Carnap according to which already in 1930, when Tarski and Carnap met for the first time, "in contrast to our [the Vienna Circle's] view that there is a fundamental difference between logical and factual statements, since logical statements do not say anything about the world, Tarski maintained that the distinction was only a matter of degree" (Carnap [3], p. 30). Tarski's skepticism towards epistemically or modally loaded notions was therefore much earlier than the writing of his papers on semantics (where it can also be sensed, for example in [30], p. 420). And it is not likely that he changed his mind in the thirties (or at any other time), for in 1940–1941 we find him siding with Quine against Carnap on a famous polemic concerning the analytic-synthetic distinction. One of the things that Tarski and Quine argued against Carnap was that the notions of logical truth and of truth of classical mathematics, even if philosophically problematic, were indispensable notions that could be precisely characterized, but that the notion of analytical implication was "an unexplained notion that we were not committed hitherto," as Quine says in a letter of 1943 to Carnap that discusses the meetings of 1940–1941 (Quine [20], p. 296).

These considerations suggest that criticisms of Tarski's definition based on its inadequacy to capture (either intensionally or extensionally) some epistemically or modally loaded notion, cannot be seen as arguments for the inadequacy of Tarski's definition to accommodate Tarski's motivations. Of course, our considerations do not suggest that these criticisms, in themselves and not as criticisms of Tarski, cannot be correct. However, it is equally important to point out that Tarski's definition accommodated his main motivations, and that the concept defined precisely for a wide variety of formal languages by Tarski mirrored in very significant ways the informal logical usage of the notion of consequence concerning those same languages. It is in this way that the remarks made in this paper point to a positive justification of the adequacy of Tarski's definition. Its adequacy, that is, for the task of characterizing a certain pretheoretically important notion, with marked and distinctive features grounded in ordinary logical usage.

Acknowledgments I wish to thank Paul Benacerraf, John Etchemendy, Ignacio Jané, Scott Soames, and Bas van Fraassen for helpful discussions on the first version of this paper, and Stewart Shapiro and an anonymous referee for useful written comments that led to improvements of a later version.

NOTES

1. A structural rule is informally characterized by Tarski as one "in which only the external structure of sentences is involved" ([30], p. 410). Pending further clarification of the notion of a structural rule (which Tarski acknowledges to be necessary; see [30], p. 413, note 1), the ω-rule that Tarski immediately is going to consider would appear to satisfy the informal characterization. But probably Tarski is including other typical characteristics of the ordinary rules, such as their finitary nature, as part of the intended meaning of 'structural'.

2. For a more elaborate discussion of this passage than the one provided here, see my [13].

3. The perplexity caused by Tarski's remark about the intuitive logical validity of the ω-rule (together with the fact that its formalization in first-order arithmetic is not model-theoretically valid) appears also in Sher [22], in the context of a more sympathetic reading of Tarski than Etchemendy's. According to Sher, "the sentence 'For every natural number *n*, *Pn*' seems to follow, in some important sense, from the set of sentences '*Pn*', where *n* is a natural number, but there is no way to express this fact by the proof method for standard first-order logic. This situation, Tarski says, shows that proof theory by itself cannot fully accomplish the task of logic" (Sher [22], p. 38). However, all the sense she seems to make of Tarski's remark is that "although the relation between the set of

sentences 'Pn' and the universal quantification '$(\forall x)\,Px$', where x ranges over the natural numbers and 'n' stands for a name of a natural number, is not logical consequence, we will be able to characterize it accurately within the framework of Tarskian semantics, e.g., in terms of ω-completeness" (Sher [22], p. 39).

4. It is not clear what they are trickier than, since A might itself be a "Gödel sentence"; in fact, Gödel proved in [12] the ω-incompleteness of a certain formalization of arithmetic, for the "Gödel sentence" is a sentence of the form of A all of whose numeral instances are provable in Gödel's formalization of Peano arithmetic. Etchemendy is perhaps making a distinction between undecidable sentences of the form of A and other undecidable sentences.

5. It seems clear that Gödel proved his results for a system with the arithmetical symbols as primitives for reasons of simplicity (among perhaps others), but a great part of their significance at the time (especially for philosophers sympathizing with logicism) was based on the fact that they could be easily seen to apply to the theory of types of *Principia Mathematica*, which was the logical framework of reference at the time (as well as to other theories, more powerful mathematically, in which arithmetic could be developed, like Zermelo's set theory).

6. The converse is false, for there are nondenumerably many such invariant notions, but only denumerably many are definable in the theory of types. This is surely the explanation for Tarski's use of 'logical notion' instead of 'logical term' or 'logical constant'. But of course, a logical term can be readily defined as a term (or, perhaps, a primitive term) which defines a logical notion.

7. This will be so regardless of the cardinality of the universe of individuals, even though only if this universe is infinite will the class of logically defined "numbers" be infinite (hence the need for the axiom of infinity in the logicist formalization of arithmetic). If the universe of individuals is finite, the class of logically defined "numbers" will be finite, but it will still be the class denoted by the logically defined predicate 'to be a natural number'.

8. That no apparent mention of domain variation is made in Tarski's definition was also emphasized in some papers by Corcoran (see [5], p. 43 and [6], p. 70).

9. The history of how first-order logic became isolated as a useful and interesting fragment of more comprehensive and powerful logical languages is documented in Moore [19].

10. The example Tarski has in mind involves quite a few complications, and we can describe it only superficially. The example is the "calculus of classes," for which most of the results of Tarski [27] are proved. This is syntactically a first-order language without identity and with a binary predicate interpreted intuitively as the predicate of inclusion between classes. The theory includes a standard calculus for first-order logic and a set of postulates for inclusion given for the first time by Huntington. According to Tarski, "the calculus of classes is a fragment of mathematical logic" ([27], p. 168). The predicate of inclusion is taken as a logical constant very much like identity; its extension in a particular interpretation of the language is determined simply by the domain of individuals of the interpretation, which determines in turn its domain of classes. However, the variables of the language range over the domain of classes of the interpretation, so the semantics of the language is not a standard first-order semantics; this is possible only because the concept of class is "logical," in the sense that any domain is taken to determine univocally a set of classes (of individuals in that domain).

    There is a sentence of the language of the calculus of classes thus formalized which is not provable in the calculus, and which intuitively says that every nonempty class includes a class of one element. But this sentence is true in every individual domain. It turns out that, when it is added as an axiom, the set of provable sentences becomes the same as that of the sentences true in every individual domain. In the language of the calculus of classes there are sentences which express facts about the cardinality of the domain of individuals (and which are not provable in the theory); thus, the sentence '$\exists x \exists y \exists z (x \subseteq y \,\&\, \neg y \subseteq x \,\&\, y \subseteq z \,\&\, \neg z \subseteq y)$' is true in a domain d if and only if d contains at least two elements, and simply true if the number of ("logical") individuals is at least two.

11. In fact, a more accurate translation of the German text of this footnote probably would be: "Since we accept the axiom of infinity, we of course give up the postulate according to which only the sentences which are correct in every individual domain must be provable sentences of logic."

12. In two historical surveys of early model theory and Tarski's work in model theory, Vaught, who was a close associate of Tarski, points out that the series of works by Tarski on semantical concepts (definability, truth, logical consequence), which culminates in [27] and [30], sprang from Tarski's unhappiness about the imprecise available ways for stating the results proved in the 1927–1929 seminar (see Vaught [36], pp. 160ff. and [37], pp. 870ff.).

13. The first theory is the first-order theory of the ordinal type ω with a sign for the natural relation of order among ordinals; the second, "Presburger arithmetic," the first-order theory of the whole numbers with a sign for the number one as distinguished object, and a sign for the relation of addition. On these examples, and for other references to Tarski's early work on the model theory of some first-order theories, see Tarski [29], §5 and appendix.

14. The point that the Löwenheim-Skolem-Tarski theorem presents problems for Etchemendy's view on domain variation is also stressed in Sher [22], p. 41.

15. A similar conjecture is made, without historical support, in Hodges [18], p. 138.

16. The development of Carnap's ideas on the matter in reaction to Gödel's results is documented in Coffa [4], especially pp. 285ff.

17. Etchemendy, however, has made the claim that Tarski tried to argue that the *defined* concept of logical consequence had certain modal (or epistemic) properties, and that, in the course of this argument, he committed a simple modal fallacy (see Etchemendy [10], pp. 85ff.). This claim of Etchemendy is based again on an incorrect reading of Tarski's relevant texts; I have pointed this out elsewhere (cf. [13]).

REFERENCES

[1] Carnap, R., *Logische Syntax der Sprache*, Springer, Vienna, 1934. Zbl 0182.00301 MR 38:24 4.2, 5, 5, 5, 5

[2] Carnap, R., *The Logical Syntax of Language*, Routledge & Kegan Paul, London, 1937. 4.2, 5

[3] Carnap, R., "Intellectual autobiography," pp. 1–84 in *The Philosophy of Rudolf Carnap*, edited by P. A. Schilpp, Open Court, La Salle, 1963. 5

[4] Coffa, J. A., *The Semantic Tradition from Kant to Carnap*, Cambridge University Press, Cambridge, 1991. 5

[5] Corcoran, J., "Conceptual structure of classical logic," *Philosophy and Phenomenological Research*, vol. 33 (1972), pp. 25–47. Zbl 0657.03005 5

[6] Corcoran, J., "Meanings of implication," *Diálogos*, vol. 4 (1973), pp. 59–76. MR 1260674 5

[7] Etchemendy, J., "The doctrine of logic as form," *Linguistics and Philosophy*, vol. 6 (1983), pp. 319–334. 1

[8] Etchemendy, J., "Tarski on truth and logical consequence," *The Journal of Symbolic Logic*, vol. 53 (1988), pp. 51–79. Zbl 0645.03002 MR 89g:01051 1, 3.1, 3.1, 3.1, 4.1

[9] Etchemendy, J., "Models, semantics and logical truth," *Linguistics and Philosophy*, vol. 11 (1988), pp. 91–106. Zbl 0653.03004 1

[10] Etchemendy, J., *The Concept of Logical Consequence*, Harvard University Press, Cambridge, 1990. Zbl 0743.03002 MR 92f:03003 1, 3.1, 3.1, 5

[11] Gödel, K., "Die Vollständigkeit der Axiome des logischen Funktionenkalküls," *Monatshefte für Mathematik und Physik*, vol. 37 (1930), pp. 349–360. Translated as "The completeness of the axioms of the functional calculus of logic," pp. 583–591 in *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, edited by J. van Heijenoort, Harvard University Press, Cambridge, 1967. 4.2

[12] Gödel, K., "Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I," *Monatshefte für Mathematik und Physik*, vol. 38 (1931), pp. 173–198. Translated as "On formally undecidable propositions of *Principia Mathematica* and Related Systems I," pp. 596–616 in *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, edited by J. van Heijenoort, Harvard University Press, Cambridge, 1967. 3.3, 3.3, 4.3, 5

[13] Gómez-Torrente, M., "On a fallacy attributed to Tarski," forthcoming in *History and Philosophy of Logic*. Zbl 01678958 MR 2000f:03006 5, 5

[14] Hempel, C. G., "On the nature of mathematical truth," pp. 377–393 in *The Philosophy of Mathematics*, second edition, edited by P. Benacerraf and H. Putnam, Cambridge University Press, Cambridge, 1983. Zbl 0060.01909 MR 7,193b 4.2

[15] Hilbert, D., and W. Ackermann, *Grundzüge der theoretischen Logik*, Springer, Berlin, 1928. 4.1, 4.2, 4.2

[16] Hilbert, D., and W. Ackermann, *Grundzüge der theoretischen Logik*, second edition, Springer, Berlin, 1938. Zbl 0018.19301 3.2, 3.2

[17] Hilbert, D., and W. Ackermann, *Principles of Mathematical Logic*, Chelsea, New York, 1950. Zbl 0040.00402 3.2

[18] Hodges, W., "Truth in a structure," *Proceedings of the Aristotelian Society*, new series, vol. 86 (1985–86), pp. 135–151. 5

[19] Moore, G. H., "The emergence of first-order logic," pp. 95–135 in *History and Philosophy of Modern Mathematics*, edited by W. Aspray and P. Kitcher, Minnesota Studies in the Philosophy of Science, vol. XI, University of Minnesota Press, Minneapolis, 1988. MR 89i:03001 5

[20] Quine, W. V., "Quine to Carnap," pp. 294–300 in *Dear Carnap, Dear Van. The Quine-Carnap Correspondence and Related Work*, University of California Press, Berkeley, 1990. 5

[21] Ramsey, F. P., "The foundations of mathematics," pp. 1–61 in *The Foundations of Mathematics and other Logical Essays*, Kegan Paul, London, 1931. 4.2

[22] Sher, G., *The Bounds of Logic*, M.I.T. Press, Cambridge, 1991. MR 94g:03014 5, 5, 5, 5

[23] Tarski, A., "On definable sets of real numbers," pp. 110–142 in *Logic, Semantics, Metamathematics*, second edition, Hackett, Indianapolis, 1983. 4.3

[24] Tarski, A., "Some observations on the concepts of ω-consistency and ω-completeness," pp. 279–295 in *Logic, Semantics, Metamathematics*, second edition, Hackett, Indianapolis, 1983. 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 4.2, 4.2

[25] Tarski, A., "Bemerkung der Redaktion," *Fundamenta Mathematicæ*, vol. 23 (1934), p. 161. 4.3, 4.3

[26] Tarski, A., and A. Lindenbaum, "On the limitations of the means of expression of deductive theories," pp. 384–392 in *Logic, Semantics, Metamathematics*, second edition, Hackett, Indianapolis, 1983. 3.3, 4.3

[27] Tarski, A., "The concept of truth in formalized languages," pp. 152–278 in *Logic, Semantics, Metamathematics*, second edition, Hackett, Indianapolis, 1983. 3.2, 3.2, 3.2, 3.2, 3.3, 4.1, 4.2, 4.2, 4.2, 4.2, 4.2, 4.2, 5, 5, 5

[28] Tarski, A., "On the foundations of Boolean algebra," pp. 320–341 in *Logic, Semantics, Metamathematics*, second edition, Hackett, Indianapolis, 1983. 4.3

[29] Tarski, A., "Foundations of the calculus of systems," pp. 343–383 in *Logic, Semantics, Metamathematics*, second edition, Hackett, Indianapolis, 1983. 2, 4.3, 5

[30] Tarski, A., "On the concept of logical consequence," pp. 409–420 in *Logic, Semantics, Metamathematics*, second edition, Hackett, Indianapolis, 1983. 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3.2, 3.2, 3.2, 3.2, 3.3, 3.3, 4.1, 4.1, 4.1, 4.1, 4.3, 4.3, 4.3, 4.3, 5, 5, 5, 5, 5, 5

[31] Tarski, A., *Einführung in die mathematische Logik und in die Methodologie der Mathematik*, Springer, Vienna, 1937. Zbl 0018.00101 3.2, 3.3, 4.3, 4.3, 4.3, 4.3, 4.3, 4.3, 4.3

[32] Tarski, A., *Introduction to Logic and to the Methodology of Deductive Sciences*, Oxford University Press, New York, 1941. Zbl 0025.00401 MR 2,209a 3.2, 3.3, 4.3, 4.3, 4.3, 4.3, 4.3

[33] Tarski, A., "A philosophical letter of Alfred Tarski," *Journal of Philosophy*, vol. 84 (1987), pp. 28–32. MR 88d:03008 3.3

[34] Tarski, A., "What are logical notions?," *History and Philosophy of Logic*, vol. 7 (1986), pp. 143–154. Zbl 0622.03004 MR 88b:03010 3.3, 3.3, 3.3

[35] Tarski, A., "A general method in proofs of undecidability," pp. 1–35 in *Undecidable Theories*, by A. Tarski, A. Mostowski, and R. Robinson, North-Holland, Amsterdam, 1953. 4.1, 4.1, 4.1

[36] Vaught, R. L., "Model theory before 1945," pp. 153–172 in *Proceedings of the Tarski Symposium*, edited by L. Henkin et al., *Proceedings of Symposia in Pure Mathematics*, vol. XXV, American Mathematical Society, Providence, 1974. Zbl 0306.02005 MR 57:12157 5

[37] Vaught, R. L., "Tarski's work in model theory," *The Journal of Symbolic Logic*, vol. 51 (1986), pp. 869–882. Zbl 0619.03007 MR 88g:03008a 5

Department of Philosophy
Princeton University
Princeton, New Jersey 08544
email: mariog@princeton.edu