Tarski's Conceptual Analysis of Semantical Notions [1]

Solomon Feferman

Dedicated to the memory of Robert L. Vaught (1926-2002)
Fellow student, dear friend, colleague

Abstract:

Tarski is famous for his widely accepted conceptual analysis (or, in his terms, "explication") of the notion of truth for formal languages and the allied notions of satisfaction, definability, and logical consequence. From an historical point of view, two questions are of interest. First, what motivated Tarski to make these analyses, and second, what led to their particular form? The latter question is easy to answer at one level: Tarski was heavily influenced by the visible success of conceptual analysis in set-theoretic topology as practiced by the leading mathematicians at the University of Warsaw in the 1920s, and so formulated his analyses of semantical concepts in general set-theoretical terms. But the actual forms which his definitions took are puzzling in a number of respects. The question of motivation is also difficult because there was no prima facie compelling reason for dealing in precise terms with the semantical notions. These had been used quite confidently, without any such explication, by a number of Tarski's contemporaries, including Skolem and Gödel. The aim of this paper is to throw greater light on both the "why" and "how" questions concerning Tarski's conceptual analysis of semantical notions, especially that of truth.

The puzzles of "why" and "how". The two most famous and--in the view of many--most important examples of conceptual analysis in twentieth century logic were Alfred Tarski's definition of *truth* and Alan Turing's definition of *computability*. In both cases a prior, extensively used, informal or intuitive concept was replaced by one defined in precise mathematical terms. It is of historical, mathematical and philosophical interest in each such case of conceptual analysis to find out *why* and *how* that analysis was undertaken. That is, to what need did it respond, and in what terms was the analysis given? In the case of Turing, the "how" part was convincingly provided in terms of the general notion of a computing machine, and one can give a give a one-line answer to the "why" part of the question. Namely, a precise notion of computability was needed to show

[1] Expanded text of a lecture for the Colloque, "Sémantique et épistémologie", Casablanca, April 24-26, 2002, sponsored by the Fondation du Roi Abdul-Aziz al Saoud pour les Etudes Islamiques et les Sciences Humaines. I wish to thank the Fondation for its generosity in making that stimulating Colloque possible, as well as its organizers and, in particular, Dr. Hourya Benis Sinaceur for the invitation to take part in it.

that certain problems (and specifically the *Entscheidungsproblem* in logic) are *uncomputable*; prior to that, the informal concept of computability sufficed for all positive applications. I shall argue that there was no similarly compelling *logical* reason for Tarski's work on the concept of truth, and will suggest instead a combination of *psychological* and *programmatic* reasons. On the other hand, the "how" part in Tarski's case at one level receives a simple one-line answer: his definition of truth is given in general set-theoretical terms. That also characterizes his analyses of the semantical concepts of *definability, logical consequence* and *logical operation.* In fact, all of Tarski's work in logic and mathematics is distinguished by its resolute employment of set-theoretical concepts. However, the form in which these were employed shifted over time and the relations between the different accounts are in some respects rather puzzling. It is my aim here to educe from the available evidence the nature and reasons for these shifts and thereby to throw greater light on both the "why" and "how" questions concerning Tarski's conceptual analyses of semantical notions, especially that of truth. The main puzzle to be dealt with has to do with the relations between the notions of truth in a structure and absolute truth.[2]

The influence of the set-theoretical topologists. A year ago, I gave a lecture entitled "Tarski's conception of logic" for the Tarski Centenary Conference held in Warsaw at the end of May 2001 (to appear as 2002). In that I emphasized several points relevant to the questions I'm addressing here, and will take the liberty in this section of repeating the following one almost verbatim. Namely, I traced Tarski's set-theoretic approach to conceptual analysis back to his mathematical studies at the University of Warsaw during the years 1919-1924, alongside his logical studies with Stanislaw Lesniewski and Jan Lukasiewicz. Tarski's choice of concentration on mathematics and logic in this period was fortuitous due to the phenomenal intellectual explosion in these subjects in Poland following its independence in 1918. On the side of logic this has been richly detailed by Jan Wolenski in his indispensable book about the Lvov-Warsaw school (1989). A valuable account on the mathematical side is given in the little volume of "remembrances and reflections" by Kazimierz Kuratowski, *A Half Century of Polish Mathematics* (1980). The grounds for the post-war explosion in Polish mathematics were laid by a

[2] Increasing attention has been given in recent years to the shifts and puzzles in Tarski's work on conceptual analysis. Most useful to me here have been the articles of Hodges (1985/1986) and (2002), Gómez-Torrente (1996), de Rouilhan (1998), Sinaceur (2001), and Sundholm (2002), and the book of Ferreiros (1999), especially pp. 350-356. Givant (1999) provides a very clear account of the progression of Tarski's work over his entire career.

young professor, Zygmunt Janiszewski. He had obtained a doctor's degree in the then newly developing subject of topology in Paris in 1912, and was appointed, along with the topologist, Stefan Mazurkiewicz, to the faculty of mathematics at the University of Warsaw in 1915. It was Janiszewski's brilliant idea to establish a distinctive Polish school of mathematics and to make an impact on the international scene by founding a new journal called *Fundamenta Mathematicae* devoted entirely to a few subjects undergoing active development.[3] Namely, it was to concentrate on the modern directions of set theory, topology, mathematical logic and the foundations of mathematics that had begun to flourish in Western Europe early in the Twentieth Century.

Tarski's teachers in mathematics at the University of Warsaw were the young and vital Waclaw Sierpinski, Stefan Mazurkiewicz and Kazimierz Kuratowski; Sierpinski and Mazurkiewicz were professors and Kuratowski was a docent. In 1919, the year that Tarski began his studies, the old man of the group was Sierpinski, aged thirty-seven; Mazurkiewicz was thirty-one, while Kuratowski at twenty-three was the "baby". The senior member in the Warsaw mathematics department, Waclaw Sierpinski, was especially noted for his work in set theory, a subject that Tarski took up with a vengeance directly following his doctoral work on Lesniewski's system of protothetic. Though Cantorian set theory was still greeted in some quarters with much suspicion and hostility, it was due to such people as Sierpinski in Poland and Hausdorff in Germany that it was transformed into a systematic field that could be pursued with as much confidence as more traditional parts of mathematics.

The main thing relevant to the present subject that I emphasized in my Warsaw lecture concerning this background in Tarski's studies is that in the 1920s, the period of his intellectual maturation in mathematics, topology was dominated by the set-theoretical approach, and its great progress lay as much in conceptual analysis as in new results. We take the definitions of the concepts of limit point, closed set, open set, connected set, compact set, continuous function, and homeomorphism--to name only some of the most basic ones--so much for granted that it takes some effort to put ourselves back in the frame of mind of that fast-evolving era in which such definitions were formulated and came to be accepted. Of course, some of the ideas of general topology go back to Cantor and Weierstrass, but it was not until the 1910s that it emerged as a subject in its own right. Tarski couldn't have missed being impressed by the evident success of that work in its use of general set theory in turning vague informal concepts into precise definitions, in terms of which definite and often remarkable theorems could be proved.

[3] Sadly, Janiszewski died in the flu epidemic of 1919-20 and did not live to see its first issue.

The paradigmatic case of dimension. In particular, a very interesting case of conceptual analysis in topology took place during Tarski's student days. This concerned the idea of the *dimension* of various geometrical objects and began with a puzzle over how to define in precise terms what it means to be a *curve* as a one-dimensional set. Informally, the idea of a curve that had been used up to the 1800s was that of a figure traced out by a moving point. As part of the progressive rigorization of analysis in the nineteenth century, Camille Jordan had proposed to define a curve (for example, in the plane or in space) as the continuous image of a line segment. When Giuseppe Peano showed, quite surprisingly, that the continuous image of a line segment could fill up a square in the plane, a new definition was urgently called for. There were a number of candidates for that, but the one relevant to our story and one that succeeded where Jordan's definition failed is that provided by Karl Menger in Vienna in 1921.[4] The details of his definition, which explains in quite general topological terms which sets in a topological space can be assigned a natural number as dimension, are not important for the present story. What *is* important is that Menger was soon in communication with the Warsaw topologists, including Kuratowski and Bronislaw Knaster (who was a close friend of Tarski), and that his conceptual analysis of the notion of dimension had a direct impact on their work. In Tarski's 1931 paper on definable sets of real numbers, the notion of dimension (among other intuitive geometrical notions) is specifically referred to as a successful example of conceptual analysis.[5] My conclusion is that for Tarski, topology was paradigmatic in its use of set theory for conceptual analysis.

In my Warsaw lecture I went in some detail into the *form* of Tarski's use of set theory in his analyses of the concepts of truth, logical consequence and of what is a logical notion, and how it was that in this last, Tarski assimilated logic to higher set theory to (what I regard as) an unjustified extent.[6] But that only answers the "how" part of our basic question at one level, and for a fuller answer, one must probe deeper in each case. That will only be done here for the concept of truth; for the case of logical consequence, see the rewarding discussion in Gómez-Torrente (1996).

[4] See Menger (1994), pp. 38ff. An essentially equivalent definition was given by Paul Urysohn in Moscow, independently of Menger's and around the same time.

[5] See p. 112 of the English translation of (1931) in Tarski (1983). Interestingly, he remarks loc. cit. that, in contrast to the geometrical examples--in which there are competing conceptual analyses in mathematical terms because the informal notions are a confused mix--the "arbitrariness" in that of definability and related logical notions is "reduced almost to zero" because the intuitions to which they respond "are more clear and conscious."

[6] For my critique of Tarski's analysis of what is a logical notion, see Feferman (1999).

When did Tarski define truth in a structure? In the case of truth, the first puzzle has to do with the *prima facie* discrepancy between what logicians nowadays usually say Tarski did, and what one finds in the *Wahrheitsbegriff*. This was brought out in the very perspicacious paper by Wilfrid Hodges, "Truth in a structure" (1985/86), which begins with an informal explanation of the current conception. Hodges then goes on to report (p. 137) that:

> [a] few years ago I had a disconcerting experience. I read Tarski's famous monograph 'The concept of truth in formalized languages' (1935) to see what he says himself about the notion of truth in a structure. The notion was simply not there. This seemed curious, so I looked in other papers of Tarski. As far as I could discover, the notion first appears in Tarski's address (1952) to the 1950 International Congress of Mathematicians, and his paper 'Contributions to the theory of models I' (1954). But even in those papers he doesn't define it. In the first paper he mentions the notion only in order to explain that he won't be needing it for the purpose in hand. In the second paper he simply says "We assume it to be clear under what conditions a sentence ... is *satisfied* in a system..."

Hodges continues, "I believe that the first time Tarski explicitly presented the mathematical definition of truth in a structure was his joint paper (1957) with Robert Vaught." In fact, the general notion of structure for a first-order language L is already described in Tarski's 1950 ICM address, essentially as follows: a structure $\underline{A}$ is a sequence consisting of a non-empty domain A of objects together with an assignment to each basic relation, operation and constant symbol of L of a corresponding relation between elements of A, operation on elements of A, or member of A, resp.. Moreover, while Hodges is correct in saying that the notion of truth in a structure is not defined there, Tarski does talk of the antecedent notion of satisfaction in a structure as if it is well understood, since he refers to the association with each formula φ of the set of all sequences from A which satisfy φ in $\underline{A}$.

At any rate, what *is* of interest, as Hodges makes clear, is that these notions of structure, and of satisfaction and truth in a structure, do not seem to appear explicitly prior to the 1950s in Tarski's work. This is doubly puzzling, since, as a common informal notion, the idea of a structure being a model of a system of axioms well precedes that, and surely goes back to the nineteenth century. Most famously in that period, one had the stunning revelation of various models for non-Euclidean geometry. Then came Dedekind's characterization of the natural numbers and the real numbers, in both cases structurally as models, unique up to isomorphism, of suitable (second-order) axioms. That was followed by Hilbert's model-theoretic considerations concerning his axioms for geometry. The structural view of mathematics took hold in the early part of the twentieth century in algebra, analysis and topology with the formulation of axioms for groups, rings, fields, metric spaces, normed spaces, topological spaces and so on, and with the systematic exploration of their various models. Within logic, the informal concept of model has been traced back by Scanlan (1991) to the American "postulate theorists", launched by work of Huntington in 1902 and Veblen in 1904; this involved, among others the concept of consistency in the sense of satisfiability and that of categoricity.[7] Finally, it was central to the famous theorems of Löwenheim of 1915 and Skolem in 1920 on existence of countable models, and of course to Gödel's completeness theorem published in 1930.

Was Tarski a model-theorist? For Tarski, on the face of it, there were two loci of interest in the relation between axiom systems and their models.[8] The first was geometry, which began to figure in his work almost as soon as he started publishing research papers in 1924. In the year 1926-27 he lectured at Warsaw University on an elegant new axiom system for Euclidean geometry which, in distinction to Hilbert's famous system of 1899, was formulated without the use of set-theoretical notions, i.e. in first-order logic, and he considered various models of subsystems of his system in order to establish some independence results. But this work was not published until the 1960s, when Tarski had clearly shifted to the current model-theoretic way of thinking (see Tarski and Givant 1999).

The second locus was the method of elimination of quantifiers to arrive at decision procedures for all first-order statements true in certain models or classes of models. That method had been developed initially by Skolem, who applied it to the monadic theory of identity, and Langford, who applied it to the theory of dense order. The method was pursued intensively in Tarski's seminar during the years 1926-1928, beginning with his extension of Langford's work to the class of discrete orders. The most famous results of that seminar were Presburger's decision procedure for the structure of the integers with the operation of addition and the order relation, and Tarski's procedure for the real numbers with the operations of addition and multiplication, obtained by 1931.[9] The first intended

---

[7] Incidentally, Scanlan points out that Tarski was aware of that work in his abstract (1924), where a second-order system of axioms is given for the order relation on ordinals which, when restricted to the accessible ordinals, is said to be categorical "au sens de Veblen-Huntington".
[8] In addition to the two here, Hodges (personal communication) has suggested a third: propositional logics and their matrix models.
[9] See Vaught (1986) for the historical development, though presented in current model-theoretic terms.

exposition of the latter was the monograph (1967) whose scheduled publication in 1940 was postponed indefinitely because of the war. Rather than containing a model-theoretic statement of the results, this is devoted to establishing the *completeness of axiom systems* for elementary algebra and geometry, and by its methods as providing a *decision procedure for provability* in those systems. Chronologically, it is not until the full exposition of the elimination of quantifiers method in the report (1948), prepared with the assistance of J. C. C. McKinsey, that it is presented frankly as a *decision procedure for truth* in the structure of real numbers. There is a corresponding marked difference in the titles between the two publications.

Tarski's acceptance of type theory as a general framework. Hourya Sinaceur (2000, pp. 8-9) has emphasized this difference between Tarski's point of view prior to 1940 and his shift to our current way of thinking, perhaps around that time. Indeed, in the primary relevant pre-war publications, specific mathematical theories are always regarded by Tarski within an axiomatic framework, often expanded to the simple theory of types, and he refrained from speaking of structures as if they were independently existing entities. As is documented fully in Ferreirós (1999), pp. 350-356, this was the accepted way of formulating things for a number of logicians and philosophers in the 1930s, under the powerful residual influence of *Principia Mathematica*.[10] We have only to look at the title of Gödel's incompleteness paper (1931) for the most famous example, where the system actually referred to is a form of the simple theory of types based on the natural numbers as the individuals. Tarski carried so far the identification of mathematical concepts with those that can be developed in the simple theory of types, that he wrote the following in his introductory textbook on logic, after sketching how the natural numbers can be treated as classes of classes and thus based "on the laws of logic alone":

> the ...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,

[10] The shift from the ramified theory of types (RTT), as a basis for "all" of mathematics in the *Principia*, to the simple theory of types (STT) was given impetus in publications by Chwistek in 1920 and Ramsey in 1926, but was also spread informally by Carnap among others. (Reck (2002) traces the ideas for STT back to Frege's *Begriffsschrift* and to lectures that Frege gave in Jena in the early 1910s, lectures that Carnap attended.) In the well-known logic text by Hilbert and Ackermann published in 1928, analysis is formulated within RTT, but in the second edition ten years later, it is formulated in STT over the rational numbers as the individuals. Others who adopted some form of type theory to some extent or other as a general logical framework in the 1930s were Carnap, Church, Gödel, Quine and Tarski; initially, Tarski

> constitutes one of the grandest achievements of recent logical investigations. (1941, p. 81)[11]

Insofar as this statement regards the simple theory of types, necessarily with the axiom of infinity, as a part of pure logic, Tarski here blithely subscribes to the logicist program, thereby ignoring the fact that the infinity axiom is not a logical principle and that the platonist ontology normally seen to be required to justify the impredicative comprehension axioms of the theory put their logical status into question.[12] At the same time, the formulation of mathematical notions in axiomatic terms as deductive theories on their own or within a wider "logical" framework, such as the theory of types, may be related to Tarski's (later) professed nominalistic, anti-platonistic, tendencies (cf. Feferman 1999a, p. 61), but that in turn is clearly in tension with his thorough-going use of set theory in practice, and his acceptance from the beginning of Zermelo's axioms as a framework for his extensive purely set-theoretical work.[13] It should be remarked that over a long period Tarski tended to regard the simple theory of types with the axiom of infinity and Zermelo's axioms as merely alternative ways of formalizing the general theory of sets; see, for example his paper with Lindenbaum on the theory of sets (1926, p. 299) and his posthumous paper on what are logical notions (1986a, p. 151).

Yes, Tarski *was* a model-theorist. So far, we have been revolving around the side question as to why Tarski, in his primary publications prior to 1948, did not take a straightforward informal model-theoretic way of presenting various of his notions and results. But, in fact there is secondary published evidence in that period that Tarski *did* think in just those terms, including the following:
(i) in the abstract (1924) he speaks of the categoricity of a (second-order) system of axioms for the ordinals up to the first inaccessible (cf. ftn. 7 above); this work is presented in more detail in sec. 4 of Lindenbaum and Tarski (1926);

---

was also strongly influenced by Lesniewski's theory of semantical categories in this respect. For a full account of type theory at its zenith, cf. Ferreirós (1999) 350-356.

[11] The passage, translated directly from the 1937 German edition, is left unchanged through the most recent fourth English edition.

[12] One must confront this form of logicism with the equivocal statements Tarski made in his 1966 lecture "What are logical notions?", published posthumously as (1986a); see op. cit., pp. 151-153, and Feferman (1999), pp. 48-49.

[13] It also seems to me to be in disaccord with his approach to metamathematics in ordinary set-theoretical terms, in contrast to Hilbert's finitist program for metamathematics, for which formulation of mathematics in axiomatic terms was the essential point of departure.

(ii) in editorial remarks (to a 1934 paper of Skolem's in *Fundamenta Mathematicae*) concerning results obtained in the seminar led by Tarski at the University of Warsaw in 1927-28 he characterizes the models of the first-order truths of $\langle \omega, < \rangle$ (cf. (1986), vol. 4, p. 568);

(iii) in a further editorial remark (ibid.), Tarski says that he obtained an upward form of the Löwenheim-Skolem theorem in the period of the aforementioned seminar;

(iv) in the appendix to (1935/36) Tarski (informally) defines for each order type α the set $T(\alpha)$ of all elementary properties true of any pair consisting of a set X and binary relation R that orders X in order type α; he then defines two order types α and β to be *elementarily equivalent* when $T(\alpha) = T(\beta)$ and gives various examples for which this holds and for which it doesn't hold; in a footnote he says that these notions can be applied to arbitrary relations, not just ordering relations; at the end of this Appendix it is stated that these ideas emerged in the Warsaw seminar of 1926-28, but that he was able to state them in "a correct and precise form" only with the help of the methods later used to define the notion of truth;

(v) in the same Appendix, using the fact that $T(\omega) = T(\omega + \omega^* + \omega)$ Tarski concludes that the property of being a well-ordering is not expressible in first-order terms;

(vi) in an abstract with Mostowski published in the *Bulletin of the American Mathematical Society* for 1949, reporting on results obtained in 1941, a decision procedure for truths in all well-ordered systems $\langle A, < \rangle$ is said to have been established, and the relation of elementary equivalence between such systems is characterized in terms of order-types (cf. (1986) vol. 4, p. 583); by (v) there is no prior axiom system for these systems with which one is dealing.

My own conclusion from this part of the evidence is that Tarski, just like the early model-theorists who preceded him, worked comfortably with the informal notion of model for first-order and second-order languages at least since 1924. Perhaps it was only the use of type theory as the logical standard of the times that caused him to refrain from a frankly model-theoretic way of presenting his results in that area. Moreover, there were no uncertainties or anomalies in informal model-theoretic work that would have created any urgent need for conceptual analysis of semantical notions to set matters right. In Robert Vaught's valuable survey (1974) of model theory before 1945, he points out that Skolem, for example, worked comfortably with the notion of truth in a model, though--by contrast--uncertainly with the notion of proof, and the latter is a principal reason that he missed establishing the completeness theorem for first-order logic. More generally, according to Vaught (op. cit., p. 161), since the notion of truth of a first-order sentence σ in a structure A "is highly intuitive (and perfectly clear for any definite σ), it had been possible to go even as far as the completeness theorem by treating truth (consciously or unconsciously) essentially as an undefined notion--one with many obvious properties." Even Tarski, as quoted in ftn. 5 above, from (1931), agreed with that.

To what needs was Tarski responding? So *why*, then, to return to our basic question, did Tarski feel it necessary to provide an explicit definition of this notion? Vaught's own answer (op. cit., pp. 160-161), is that:

> Tarski appears to have been unhappy about various results obtained during the seminar [of 1926-28] because he felt that he did not have a precise way of stating them (see the last page of Tarski (1935/1936)). ... Tarski had become dissatisfied with the notion of truth as it was being used [informally].[14]

If anything is clear about Tarski both from all his publications and of the experience of those who worked with him, it is that he was unhappy about anything that could not be explained in precise terms, and that he took great pains in each case to develop every topic in a very systematic way from the ground up.[15] More than conceptual analysis was at issue: in his several papers of the 1930s on the methodology of deductive sciences (or calculus of systems), his aim was to organize metamathematics in quite general terms within which the familiar concepts (concerning axiomatic systems) of consistency, completeness, independence and finite axiomatizability would be explained and have their widest applicability. I was referring in part to Tarski's drives to do things in this way when I said that he was responding to *psychological* and *programmatic* needs rather than a *logical* need in the case of his conceptual analyses. But there are further psychological components involved. Namely, despite his training in logic by philosophers, Tarski was first and foremost a mathematician who specialized principally in logic, and he was first and foremost very concerned to interest mathematicians outside of logic in its concepts and results, most specifically those obtained in model theory. However, he may have thought

[14] Givant (1999) p. 52 offers related reasons for Tarski's aims in this respect.

[15] As one example, not long ago I had occasion to look back at the notes I took from Tarski's lectures on metamathematics at U. C. Berkeley in the year 1949-50. Fully a month was taken up in those lectures with developing concatenation theory from scratch before one arrived at the syntax of first-order languages.

that he could not make clear to mathematicians that these results were part of mathematics until he showed how all the logical notions involved could be defined in precise mathematical (i.e., general set-theoretical) terms.[16] On the other hand, there is no evidence that mathematicians of the time who might have been interested in the relevant model-theoretic results turned away from them as long as such definitions were lacking or worried about them for other reasons .

Definability (and truth?) in a structure--for mathematicians. What *had* worried logicians, and mathematicians more generally, in the early twentieth century was the appearance of paradoxes in the foundations of mathematics. The famous ones, such as those of Cantor, Burali-Forti, or Russell, were set-theoretical. Around the same time attention was drawn to semantical paradoxes, such as those of the Liar, or of Richard and Berry, concerning truth and definability, respectively. While these did not seem to have anything to do with questions of truth and definability in the algebraic and geometrical structures of the sort with which Tarski had been dealing, they affected him in the following way, that I consider to be a further psychological aspect of the "why" problem. Namely, from early on he seemed to think that it was the metamathematical (i.e. syntactic) form in which those concepts were defined that was a principal obstacle to mathematicians' appreciation of the subject, if not outside of the purview of mathematics altogether. Thus, for example, at the outset of his 1931 paper "On definable sets of real numbers", he writes:

> Mathematicians, in general, do not like to deal with the notion of definability; their attitude toward this notion is one of distrust and reserve. The reasons for this aversion are quite clear and understandable. To begin with, the meaning of the term 'definable' is not unambiguous: whether a given notion is definable depends on the deductive system in which it is studied ... It is thus possible to use the notion of definability only in a relative sense. This fact has often been neglected in mathematical considerations and has been the source of numerous contradictions, of which the classical example is furnished by the well-known antinomy of Richard. The distrust of mathematicians towards the notion in question is reinforced by the current opinion that this notion is outside the proper limits of mathematics altogether. The problems of making its meaning more precise, of removing the confusions and misunderstandings connected with it, and

[16] Vaught (1974) p. 161, makes a similar remark, adding, "[i]t seems clear that this whole state of affairs [of the use prior to Tarski's work of semantical notions as undefined concepts] was bound to cause a lack of sure-footedness in meta-logic." But see ftn. 5 above.

> of establishing its fundamental properties belong to another branch of science--metamathematics.[17]

Tarski goes on to say that "without doubt the notion of definability as usually conceived is of a metamathematical origin" and that he has "found a general method which allows us to construct a rigorous metamathematical definition of this notion". But then he says that

> by analyzing the definition thus obtained it proves to be possible ... to replace it by [one] formulated exclusively in mathematical terms. Under this new definition the notion of definability does not differ from other mathematical notions and need not arouse either fears or doubts; it can be discussed entirely within the domain of normal mathematical reasoning.

Technically, what Tarski is concerned with in this 1931 paper was to explain first in *metamathematical* terms and then in what he called *mathematical* terms the notion of definable sets and relations (or sets of finite sequences) in the specific case of the real numbers. More precisely, the structure in question is taken to be the real numbers with the order relation, the operation of addition, and the unit element, treated axiomatically within a form of simple type theory over that structure. The metamathematical explanation of definability is given in terms of the notion of satisfaction, whose definition is only indicated there. Under the mathematical definition, on the other hand, the definable sets and relations (of order 1 in the type structure) are simply those generated from certain primitive sets of finite sequences corresponding to the atomic formulas, by means of Boolean operations and the operations of projection and its dual.[18] Later in the paper, it is indicated how to generalize this to definability over an arbitrary structure, as introduced by the following passage:

> In order to deprive the notion of elementary definability (of order 1) of its accidental character, it is necessary to relativize it to an arbitrary system of primitive concepts or--more precisely--to an arbitrary family of primitive sets of [finite] sequences. In this relativization we no longer have in mind the primitive concepts of a certain special science, e.g. of the arithmetic of real numbers. The set *Rl* is now replaced by an arbitrary set *V* (the so-called universe of discourse or universal set) and the

[17] Quotations are from the English translation in Tarski (1983), pp.110-111.

[18] In the follow-up paper with Kuratowski, these are shown to be imbedded in the hierarchy of projective sets in Euclidean space.

symbol *Sf* is assumed to denote the set of all finite sequences *s* [of elements of] *V*; the primitive sets of sequences are certain subsets of *Sf*.[19] [20]

Since Tarski's *metamathematical* explication of the concept of *definability in a structure* makes use of satisfaction, I take it that the notion of *truth in a structure* is present implicitly in that 1931 paper. Indeed, in a footnote to the introduction he says of the metamathematical definition that "an analogous method can be successfully applied to define other concepts in the field of metamathematics, e.g., that of *true sentence* or of a *universally valid sentential function.*" Universal validity can only mean valid in every interpretation, and for that the notion of *satisfaction in a structure* is necessary. It would have been entirely natural for Tarski to spell that out for his intended mathematical audience at that stage, if he had simply regarded metamathematics as part of mathematics by presenting the syntax and semantics of first order languages as a chapter of set theory.

Truth *simpliciter*--for philosophers. Why, then, did he offer instead the puzzlingly different definition of truth that we came to know in the *Wahrheitsbegriff* (1935)*?* Actually, as Tarski makes plain in a bibliographical note to the English translation (1983, p. 152), its plan dates to 1929. According to that note, he made presentations of the leading ideas to the Logic Section of the Philosophical Society in Warsaw in October 1930 and to the Polish Philosophical Society in Lwów in December 1930. Thus it is entirely contemporaneous with the work just described on definability in a structure (and implicitly of truth in a structure), but now directed primarily to a philosophical rather than a mathematical audience.[21] Later presentations would also be similarly oriented, including his 1936 lecture to the Congrès Internationale de Philosophie Scientifique and his 1944 paper "The semantic conception of truth and the foundations of semantics" published in *Philosophy and Phenomenological Research*.

---

[19] From the translation of Tarski (1931) in (1983), p. 135.

[20] Incidentally, in my work on Hermann Weyl, I have drawn attention to his publication (1910) where a general notion of definability in a structure is proposed. Weyl's purpose was to replace Zermelo's vague notion of definite property in his axiomatization of set theory. (See van Heijenoort (1967) p. 285.) For some strange reason this paper never came to the attention of logicians such as Tarski concerned with semantical notions.

[21] This must be amplified slightly: though the *Wahrheitsbegriff* was, as well, directed to a philosophical audience via its publication in *Studia Philosophica,* the first announcement of its ideas and results was published in 1932 in the mathematical and physical sciences section of the Viennese Academy of Sciences, and the 1933 Polish monograph (of which the *Wahrheitsbegriff* was a translation), was published by the

Clearly, Tarski thought that as a side result of his work on definability and truth in a structure, he had something important to tell the philosophers that would straighten them out about the troublesome semantic paradoxes such as the Liar, by locating for them the source of those problems. Namely, on Tarski's view, everyday language is inherently inconsistent via ordinary reasoning about truth as if it were applicable to all sentences of the language. The notion of truth can be applied without contradiction only to restricted formalized languages of a certain kind, and the definition of truth for such languages requires means not expressible in the languages themselves. This, of course, makes truth prima-facie into a *relative* notion, namely relative to a language, as definability was emphasized above to be a notion relative to a structure. Nevertheless, as presented in the *Wahrheitsbegriff*, it should in my view be considered to be an *absolute* notion, albeit a fragment of such. How can that be? The difference is that we are not talking about *truth in a structure* but about *truth simpliciter*, as would be appropriate for a philosophical discussion, at least of the traditional kind. This is borne out by a number of passages, of which those that follow are only a sample (quoted from the English translation in (1983)).

After explaining the need to restrict to formalized languages of a special kind to avoid the paradoxes, Tarski writes:

> ...we are not interested here in "formal" languages and sciences in one special sense of the word "formal", namely sciences to the signs and expressions of which no meaning is attached. For such sciences the problem here discussed [of defining truth] has no relevance, it is not even meaningful. We shall always ascribe quite concrete and, for us, intelligible meanings to the signs which occur in the languages we shall consider. (1983, pp.166-7)

The definition of truth is illustrated in sec. 3 of the *Wahrheitsbegriff* for the language of the calculus of classes of the domain of individuals within the simple theory of types. Nothing is said about the nature of that domain except that it must be assumed to be infinite (cf. p. 174, ftn. 2 and p. 185); we may presume it to contain all concrete individuals. The variables of the language of the calculus of classes are then interpreted to range over arbitrary subclasses of the domain of individuals. Now, by way of contrast with the notion that he is after, Tarski takes a bow in the direction of a relative notion of truth in the following passage:

---

corresponding section of the Warsaw Academy of Sciences; see the items [32] (p.917) and [33^{m}] (p. 932) of the bibliography (Givant 1986) of Tarski's works.

> In the investigations which are in progress at the present day in the methodology of the deductive sciences (in particular in the work of the Göttingen school grouped around Hilbert) another concept of a relative character plays a much geater part than the absolute concept of truth and includes it as a special case. This is the concept of *correct* or *true sentence in an individual domain a.* By this is meant ... every sentence which would be true in the usual sense if we restricted the extension of the individuals considered to a given class *a*, or--somewhat more precisely--if we agreed to interpret the terms 'individual', 'class of individuals', etc. as 'element of the class a', 'subclass of the class a', etc., respectively.[22] (1983, p. 199)

And, further on in this connection, we have:

> ...the general concept of correct sentence in a given domain plays a great part in present day methodological researches. But it must be added that this only concerns researches whose object is mathematical logic and its parts. ... 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... (1983, pp. 239-240)

It's clear from these quotations that what Tarski is after in the *Wahrheitsbegriff* is an absolute concept of truth, relativized only in the sense that it is considered for various specific formalized languages of a restricted kind. There is of course then the question of how the meanings of the basic notions of such languages are supposed to be determined. It is also part of Tarski's project to "not make use of any semantical concept if I am not able previously to reduce it to other concepts." (1983, p. 153) Thus, meanings *can't* be given by assignments of some sort or other to the external world. Tarski's solution to this problem is to specify meanings by translations into an informally specified metalanguage associated with the given language, within which meanings are supposed to be already understood.[23] (1983, pp. 170-171)

---

[22] In a footnote, Tarski warns off the philosophical part of his audience from the relativized definition, saying that "it is not necessary for the understanding of the main theme of this work and can be omitted by those readers who are not interested in special studies in the domain of the methodology of the deductive sciences... ."

[23] Hartry Field (1972) has emphasized this as Tarski's way of solving the problem of supplying meaning, and argued that it is inadequate for a physicalist theory of truth. It is a separate, and debatable issue,

Incidentally, and this is a separate issue worthy of discussion but not pursued here, these languages are not only taken to be interpreted, but are also supposed to carry a deductive structure specified by axioms and rules of inference. Tarski requires that "[t]he sentences which are distinguished as axioms [should] seem to us to be materially true, and in choosing rules of inference we are always guided by the principle that when such rules are applied to true sentences the sentences obtained by their use should also be true."[24] (1983, p. 167) His purpose in including the deductive structure is to show what light the notion of truth throws on that of provability, but of course that is unnecessary to the task of defining truth for a given language given solely by the assumed meaning of its basic notions and the syntactic structure of its sentences.

Was Tarski a logical universalist? In a first draft of this article I argued that the distinction between treating truth in an absolute rather than relative sense has to do with that first elicited by Jean van Heijenoort in his short but innovative article, "Logic as calculus and logic as language" (1967a, 1985).[25] In brief, according to van Heijenoort, for Frege and Russell logic is a *universal language*, while the idea of logic as calculus is the approach taken in the pre-Fregean work of Boole, De Morgan and Schröder, later taken up again in the post-Russellian work beginning with Löwenheim and Skolem. To quote van Heijenoort,

> Boole has his universal class, and De Morgan his universe of discourse, denoted by '1'. But these have hardly any ontological import. They can be changed at will. The universe of discourse comprehends only what we agree to consider at a certain time, in a certain context. For Frege it cannot be a question of changing universes. One could not even say that he restricts himself to *one* universe. His universe is *the* universe. Not necessarily the physical universe, of course, because for Frege some objects are not physical. Frege's universe consists of all that there is, and it is fixed. (1985, pp. 12-13)

Russell's adaptation of this was in the ramified theory of types. The Frege-Russell viewpoint is certainly understandable if their systems are regarded as embodying purely

---

whether Tarski's program to establish semantics on a scientific basis, as described in his (1936), would require him to meet Field's demands for such a theory.

[24] In the German of (1935), the first part of this reads: "Aussagen, die als Axiome ausgezeichnet wurden, scheinen uns inhaltlich wahr zu sein..."

[25] See also van Heijenoort (1976) and Hintikka (1996).

logical notions. The work of the *Wahrheitsbegriff,* I argued, is presented in that universalist tradition, though the framework is modified to that of the pure simple theory of types (STT) rather than the ramified one, and is used informally rather than formally; in addition, as we have already remarked, the axiom of infinity is assumed in order to make use of the natural numbers within the theory. All that is in apparent conflict with the change of perspective represented by the Postscript to the *Wahrheitsbegriff,* as detailed in the very persuasive article on Tarski and the universalism of logic by Philippe de Rouilhan (1998), which was brought to my attention in the meantime.[26] On reconsideration, I have to agree that my claim needs to be qualified, though not necessarily radically; as this, too, is a side issue, but one that I want to address here, I will be as brief as the matter allows.

In the body of the *Wahrheitsbegriff* (cf. especially pp. 215 ff of (1983)), Tarski subscribed to Lesniewski's theory of semantical categories (credited to Husserl in its origins). Considered formally, this contains STT, called by Tarski in some places the theory of sets (e.g., p. 210, ftn. 2) and elsewhere the general theory of classes (e.g., pp. 241-242). Part of the significance of the theory of semantical categories is supposed to be its universality:

> The language of a complete system of logic should contain--actually or potentially--all possible semantical categories which occur in the languages of the deductive sciences. Just this fact gives to the language mentioned a certain "universal" character, and it is one of the factors to which logic owes its fundamental importance for the whole of deductive knowledge. (1983, p. 220).

Every semantical category can be assigned a natural number as order, the order of expressions of that category: the order of individual terms is lowest and the order of a relational expression is the supremum of the orders of its arguments plus one. The order of a language consonant with the doctrine of semantical categories is the supremum of the orders of the expressions in that language, thus either a natural number or the first infinite ordinal ω. A metalanguage in which truth is to be defined for a given language must be of higher order than the order of the language. In sec. 4 of the *Wahrheitsbegriff,* Tarski sketched how to define truth for languages of finite order, while in sec. 5 he argued that there is no way to do that for languages of infinite order, since there is no place for a metalanguage to go, if it is to be part of the universal language.

---

[26] See also Sundholm (2002), which is of interest as well for its account of the relations between Tarski and Lesniewski.

But in the Postscript to the *Wahrheitsbegriff*, Tarski abandoned the theory of semantical categories, so as to allow for languages of transfinite order in some sense or other. The nature of such is only sketched there and the details are problematic; the difficulties are well explained by de Rouilhan in the article mentioned above. In any case, the idea of a universal language is clearly abandoned in the Postscript, and in that sense, Tarski is *not* a universalist. But if one reformulates van Heijenoort's basic distinction in the way that Sundholm (2002) does, as being one between *languages [used] with meaning vs. languages without use [i.e. as objects of metamathematical study]*, in my view we still find Tarski positioned on the former side of the 'versus', *even* in the Postscript. How then is it that Sundholm places him on the latter side? The choice of position depends on whether one emphasizes, as I do here, that *qua* philosopher, in the *Wahrheitsbegriff* Tarski is after the concept of absolute truth, or as Sundholm does, that *qua* (meta-) mathematician, he is after a concept of truth relative to a language. And that has to do with the tension between the two sides of Tarski's efforts with respect to the semantical notions, the one represented in the *Wahrheitsbegriff* for philosophers, and the original one for mathematicians, described above.

Next we see how this opposition affects the single most famous feature of the *Wahrheitsbegriff*, the truth scheme.

What the truth scheme does and doesn't do. It is for *truth simpliciter* treated within the framework of STT that Tarski can formulate the conditions required of a "materially adequate" definition of truth for a language L of finite order in its metalanguage (inside STT). That takes the form of the scheme (or "convention", in his terminology)

(T) x is a true sentence if and only if p,

each instance of which is given by substituting for 'x' the name of a sentence in L, and for 'p' the sentence itself as it is given in its metalanguage. In particular, in the case that L is the language of a specific mathematical structure $\underline{A}$ whose underlying domain, relations, operations and distinguished elements are taken to be given in the metalanguage, we can replace 'true sentence' in the left hand side of (T) by 'sentence true in $\underline{A}$'. That is how it is done for the structure underlying the calculus of classes used to illustrate the definition of truth in the *Wahrheitsbegriff*, whose domain is the class of all subclasses of the class of individuals, and whose only relation is that of inclusion. But nothing like (T) is suggested by Tarski for the notion of truth in arbitrary structures. In particular, no analogue to the scheme (T) is suggested for the notion of truth in the calculus of classes relative to an arbitrary domain *a*. An obvious modification would take the form,

(T_{rel}) for all *a*, x is a true sentence in the domain *a* if and only if p,

each instance of which is given by substituting for ‘x’ the name of a sentence in the fragment L, and for ‘p’ the *relativization of that sentence* to the variable *a*. It is apparent from this example that if the definition of truth in a structure more generally were to be presented in the framework of STT, the formulation of a corresponding truth scheme would be all the more cumbersome, and would not have the striking obviousness of (T) as a criterion for the definition of truth. In this case, the tension between Tarski the philosopher and Tarski the (meta)-mathematician is resolved, out of mere simplicity, in favor of the former.

The philosophical impact. The actual reception by philosophers of the form of the definition of truth given to us in the *Wahrheitsbegriff* was initially mixed, and remains so to this day. Some, like Karl Popper, took to it fairly quickly; they had first met in Prague in 1934 at a conference organized by the Vienna Circle, with whose tenets regarding the nature of science Popper was in dispute. When they met again in 1935 during an extended visit that Tarski made to Vienna, Popper asked Tarski to explain his theory of truth to him:

> ... and he did so in a lecture of perhaps twenty minutes on a bench (an unforgotten bench) in the *Volksgarten* in Vienna. He also allowed me to see the sequence of proof sheets of the German translation of his great paper on the concept of truth, which were then just being sent to him... . No words can describe how much I learned from all this, and no words can express my gratitude for it. Although Tarski was only a little older than I, and although we were, in those days, on terms of considerable intimacy, I looked upon him as the one man whom I could truly regard as my teacher in philosophy. I have never learned so much from anybody else. (Popper 1974, p. 399)

Rudolf Carnap, a central figure in the Vienna Circle, was another philosopher who took reasonably quickly to Tarski’s theory of truth. He had been favorable to Tarski’s general approach to the methodology of deductive sciences since their first meeting in 1930. Later, during Tarski’s visit to Vienna in 1935, Carnap became a convert to the theory of truth and urged Tarski to present it at the forthcoming first Unity of Science conference to be held in Paris. Relating the circumstances that led him to accept and promote the theory published in Tarski's "great treatise on the concept of truth", Carnap wrote in his intellectual autobiography:

> When Tarski told me for the first time that he had constructed a definition of truth, I assumed that he had in mind a syntactical definition of logical truth or provability. I was surprised when he said that he meant truth in the customary sense, including contingent factual truth.... I recognized that [Tarski's approach] provided for the first time the means for precisely explicating many concepts used in our philosophical discussions. (Carnap 1963, p. 60)

Carnap urged Tarski to report on the concept of truth at the forthcoming congress in Paris. "I told him that all those interested in scientific philosophy and the analysis of language would welcome this new instrument with enthusiasm, and would be eager to apply it in their own philosophical work." Tarski was very skeptical. "He thought that most philosophers, even those working in modern logic, would be not only indifferent, but hostile to the explication of his semantical theory." Carnap convinced him to present it nevertheless, saying that he would emphasize the importance of semantics in his own paper, but Tarski was right to be hesitant. As Carnap reports:

> At the Congress it became clear from the reactions to the papers delivered by Tarski and myself that Tarski's skeptical predictions had been right. To my surprise, there was vehement opposition even on the side of our philosophical friends. Therefore we arranged an additional session for the discussion of this controversy outside the offical program of the Congress. There we had long and heated debates between Tarski, Mrs. Lutman-Kokoszynska, and myself on one side, and our opponents [Otto] Neurath, Arne Naess, and others on the other. (1963, p. 61)

The bone of contention was whether the semantical concepts could be reconciled with the strictly empiricist and anti-metaphysical point of view of the Vienna Circle. In the write-up (1936) of his talk "The establishment of scientific semantics" for the conference, Tarski tried to make the views compatible, but he still found it necessary to respond to critics as late as 1944 in the expository article, "The semantic conception of truth and the foundations of semantics."

I think it is fair to say that since then, at least Tarski's scheme (T) has been central to many philosophical discussions of the nature of truth,[27] though the philosophical significance of his *definition* of truth--or whether absolute truth is even definable--continues to be a matter of considerable dispute.

Semantics without syntax: one more try at a "normal" mathematical definition. In the address (1952) that he made to the 1950 International Congress of Mathematicians held in Cambridge, Massachusetts, Tarski tried once more to interest mathematicians in model theory by developing its notions in "normal" mathematical terms, i.e. without reference to the syntax of first-order languages. These notions are applied to any given similarity class of structures A, called there algebraic systems. In ordinary metamathematical terms, a subclass S of the given similarity class is said to be an *arithmetical class* (or elementary class) if for some sentence σ of the corresponding first-order language L, S consists of all structures A such that σ is true in A. As Tarski describes what he is after,

> [t]he notion of arithmetical class is of a metamathematical origin; whether or not a set [sic!] of algebraic systems is an arithmetical class depends upon the form in which its definition can be expressed. However, it has proved to be possible to characterize this notion in purely mathematical terms and to discuss it by means of normal mathematical methods. The theory of arithmetical classes has thus become a mathematical theory in the usual sense of this term, and in fact it can be regarded as a chapter of universal algebra. ((1952), p. 705, reprinted in (1986) p. 461)

The means by which this is accomplished is by a kind of uniform extension across the given similarity class of the basic relations and operations on them that Tarski had used to explain the notion of definable relation in (1931). In the terminology of (1952) these are given by *arithmetical functions* F whose domain is the given similarity class and which for each A in that class has for its value a subset of the set A^ω of infinite sequences of elements of the domain A of A. Tarski explains frankly (op. cit., pp. 706-707) that these functions are obtained by imitation of the recursive metamathematical definition of satisfaction, in terms of which each such F is determined by a formula φ of L with F(A)

[27] Cf., for example, the collections of articles in Blackburn and Simmons (1999) and in Lynch (2001).

equal to the set of all sequences which satisfy φ in A. The collection of all arithmetical functions is denoted **AF**. Certain F in **AF**, called *simple functions*, can be distinguished as corresponding to sentences, and for these, F(A) is either empty or A^{ω}. Finally, an arithmetical class S is defined to be one such that for some simple function F, A is in S if and only if F(A) = A^{ω}.

As Quine said of Russell's Axiom of Reducibility, this entire procedure is "indeed oddly devious",[28] and at a crucial point in the development, the effort at "normal" mathematization even breaks down. Namely, one of the main results of (1952) is Theorem 13, the *compactness theorem for arithmetical functions,* which takes the form that if **K** is any subclass of **AF** whose intersection is the function Z that assigns the empty set to each structure, then there is a finite subset **L** of **K** whose intersection is Z. The compactness theorem for **AC** is a corollary. Of this, Tarski says, "[a] mathematical proof of Theorem 13 is rather involved. On the other hand, this theorem easily reduces to a metamathematical result which is familiar from the literature, in fact to Gödel's completeness theorem for elementary logic." No indication is given as to what "mathematical proof" of this theorem Tarski had in mind; the first published candidate that might be considered to qualify for such would be the one using ultraproducts (a "mathematical" notion) given ten years later by Frayne, Morel and Scott (1962). But even that depends on the fundamental property of ultraproducts relating truth in such a product to truth in its factors, the formulation and proof of which makes essential use of syntax.

Though some of the language and notation such as **AC** and $\mathbf{AC}_{\delta}$ introduced in Tarski (1952) has survived in the model-theoretic literature, that of the vehicle of arithmetical functions has not, and--as far as I can tell--the impact of this approach on mathematicians outside of logic was nil.[29] Tarski himself abandoned it soon enough in favor of normal metamathematical explanations connecting semantics *to* syntax, finally fully spelled out in Tarski and Vaught (1957). Of course, all of that is unproblematically a part of ordinary set-theoretical mathematics, in accord with Tarski's basic vision of the subject.

---

[28] In van Heijenoort (1967), p. 151.

[29] And within logic it had a specific unfortunate result: at Tarski's behest, Wanda Szmielew reformulated her important elimination of quantifiers procedure for Abelian groups in terms of the language of arithmetical functions (Szmielew 1955), turning something already rather complicated syntactically into an unreadable piece that, perversely, served even further to hide the underlying mathematical facts. Eklof and Fisher (1972) subsequently reestablished her results by means of understandable standard model-theoretic techniques in a way that also brought the needed facts into relief.

Coda. In bringing his work on truth *simpliciter* to the attention of philosophers via the simply stated truth scheme (T), Tarski *did* have the kind of impact he would have liked to have had with the mathematicians in trying to interest them in semantical notions for structures via their purported mathematization. On the other hand, the enormously successful theory of models that began to take off in the 1950s, propelled by the basic work of Tarski and Abraham Robinson among others, and that has now reached applications to algebra, number theory and analysis of genuine mathematical interest, makes common use, not of its "mathematized" version of arithmetical functions and classes, but of its basic "metamathematical" version of satisfaction and truth in a structure, and has been accepted by mathematicians without qualms about those notions. So Tarski's continual concerns in that respect were, in my view, quite misplaced.

To conclude, I must return to the question implicitly raised at the beginning of this essay by the statement that--in the view of many--Tarski's definition of truth is one of the most important cases of conceptual analysis in twentieth century logic.[30] Namely, how important is it? I have been told by more than one colleague (no names, please) over the years that Tarski was merely belaboring the obvious. I have to agree that there is some justice to this criticism, at least if we're thinking about the notions of satisfaction and truth in a structure--after all, the definitions are practically forced on us. But even if that's granted, Tarski's explication of these concepts, at least in the way that it was presented in the 1950s, has proved to be important as a paradigm for all the work in recent years on the semantics of a great variety of logical and computational languages as well as parts of natural language. And it has raised interesting questions about possible other approaches to informal semantics when that paradigm doesn't seem to apply in any direct way (cf. Hodges 2002). Finally, as I have detailed in my (1999a), the definitions of satisfaction and truth have had some essential technical applications within standard metamathematics, including, besides non-definability results à la Tarski, their use in Gödel's original (and nowadays preferred) definition of constructible set and in the use of partial truth definitions for non-finite and non-bounded axiomatizability results for various theories. Though there many not have been a compelling reason for the definitions in early model theory, they now constitute a *sine qua non* of our subject.[31]

---

[30] Tarski himself thought that his two most important contributions to logic were the decision procedure for the elementary theory of real numbers and his definition of truth.

[31] I wish to thank Geoffrey Hellman, Wilfrid Hodges, Paolo Mancosu, Göran Sundholm and Johan van Benthem for their useful comments on a draft of this paper.

Departments of Mathematics and Philosophy
Stanford University

References

S. Blackburn and K. Simmons (eds.) (1999) *Truth*, Oxford Univ. Press.

R. Carnap (1963), Intellectual autobiography, in *The Philosophy of Rudolf Carnap* (P. A. Schilpp, ed.), *The Library of Living Philosophers, vol.11,* Open Court (La Salle).

P. de Roulihan (1998), Tarski et l'universalité de la logique, in *Le Formalisme en Question. Le tournant des années trente* (F. Nef and D. Vernant, eds.), Vrin (Paris), 85-102.

P. C. Eklof and E. R. Fisher (1972), The elementary theory of Abelian groups, *Annals of Mathematical Logic 4,* 115-171.

S. Feferman (1999), Logic, logics and logicism, *Notre Dame J. of Formal Logic 40* 31-54.

_________ (1999a), Tarski and Gödel between the lines, in *Alfred Tarski and the Vienna Circle* (J. Wolenski and E. Köhler, eds.), Kluwer Academic Publishers (Dordrecht), 53-63.

_________ (2002), Tarski's conception of logic; to appear in the proceedings of the Tarski Centenary Conference, Warsaw May 28-June 1, 2001.

J. Ferreirós (1999), *Labyrinth of Thought. A history of set theory and its role in modern mathematics,* Birkhäuser Verlag (Basel)

H. Field (1972), Tarski's theory of truth, *J. of Philosophy 69, no. 13,* 347-375.

T. Frayne, A. Morel and D. Scott (1962), Reduced direct products, *Fundamenta Mathematicae 51*, 195-228.

S. Givant (1986), Bibliography of Alfred Tarski, *J. of Symbolic Logic 51*, 913-941.

________ (1999), Unifying threads in Alfred Tarski's work, *The Mathematical Intelligencer 21*, 47-58.

K. Gödel (1931), Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I, *Monatshefte für Mathematik und Physik 38,* 173-198.

M. Gómez-Torrente (1996), Tarski on logical consequence, *Notre Dame Journal of Formal Logic 37,* 125-151.

L. Henkin, J. Addison, C. C. Chang, W. Craig, D. Scott, R. Vaught (1974), *Proceedings of the Tarski Symposium,* American Mathematical Society (Providence).

J. Hintikka (1996), *Lingua Universalis vs. Calculus Ratiocinator: An ultimate presupposition of twentieth-century philosophy*, Kluwer (Dordrecht).

W. Hodges (1985/1986), Truth in a structure, *Proceedings of the Aristotelian Society, new series 86*, 131-151.

_________ (2002), What languages have Tarski truth definitions?; to appear in the proceedings of the Tarski Centenary Conference, Warsaw May 28-June 1, 2001.

K. Kuratowski (1980), *A Half Century of Polish Mathematics*, Pergamon Press (Oxford).

A. Lindenbaum and A. Tarski (1926), Communication sur les recherches de la théorie des ensembles, *Comptes Rendus des Séances de la Soc. des Sciences et Lettres de Varsovie, Class III, 19*, 299-330.

M. P. Lynch (ed.) (2001), *The Nature of Truth*, MIT Press (Cambridge)

K. Menger (1994), *Reminiscences of the Vienna Circle and the Mathematical Colloquium*, Kluwer Academic Publishers (Dordrecht).

K. Popper (1974), Comments on Tarski's theory of truth, in Henkin et al. (1974), 397-409

E. Reck (2002), From Frege and Russell to Carnap: Logic and logicism in the 1920s, in *Rudolf Carnap: From Jena to L.A.* (S. Awodey and C. Klein, eds.), Open Court (Chicago), forthcoming.

M. J. Scanlan (1991), Who were the American postulate theorists?, *J. Symbolic Logic 56,* 981-1002.

H. Sinaceur (2000) Address at the Princeton University Bicentennial Conference on problems of mathematics (December 17-19, 19460, by Alfred Tarski, *Bull. Symbolic Logic 6* (1-44).

_________ (2001), Alfred Tarski: Semantic shift, heuristic shift in metamathematics, *Synthese 126*, 49-65.

G. Sundholm (2002), Languages with meaning versus languages without use, to appear.

W. Szmielew (1955), Elementary properties of Abelian groups, *Fundamenta Mathematicae 41*, 203-271.

A. Tarski (1924), Sur les principes de l'arithmétique des nombres ordinaux (transfinis), *Annales de la Société Polonaise de Mathématiques 3,* 148-149. [Reprinted in Tarski (1986), vol. 4, 533-534.]

_______ (1931), Sur les ensembles définissables de nombres réels. I, *Fundamenta Mathematicae 17*, 210-239; revised English translation in Tarski (1983),

110-142.
_______ (1935), Der Wahrheitsbegriff in den formalisierten Sprachen, *Studia Philosophica 1*, 261-405; revised English translation in Tarski (1983), 152-278.
_______ (1935/36) Grundzüge des Systemenkalküls. Erster Teil, *Fundamenta Mathematicae 25*, 503-526/ Zweiter Teil, *F.M. 26* 283-301; revised English translation as a single article in Tarski (1983), 342-383.
_______ (1936) Grundlagen der Wissenschaftlichen Semantik, *Actes du Congrès Internationale de Philosophie Scientifiques*, vol. 3, Hermann & C[ie] (Paris), 1-8; revised English translation in Tarski (1983), 401-408.
_______ (1941), *Introduction to Logic and to the Methodology of Deductive Sciences*, Oxford University Press (New York).
_______ (1944) The semantic conception of truth and the foundations of semantics, *Philosophy and Phenomenological Research 4*, 341-375.
_______ (1948), *A decision method for elementary algebra and geometry*, (prepared with the assistance of J.C.C. McKinsey), RAND Corp. (Santa Monica); second revised edition, 1951, Univ. of California Press (Berkeley).
_______ (1952), Some notions and methods on the borderline of algebra and metamathematics, *Proc. International Congress of Mathematicians, Cambridge, Mass. 1950*, Vol. 1, Amer. Math. Soc. (Providence), 705-720.
_______ (1954), Contributions to the theory of models I, *Indagationes Mathematicae 16*, 572-581.
_______ (1967), *The Completeness of Elementary Algebra and Geometry*, Inst. Blaise Pascal (Paris). (Reprint from page proofs of a work originally scheduled for publication in 1940 in the series, *Actualités Scientifiques et Industrielles*, Hermann & C[ie], Paris, but which did not appear due to the wartime conditions.)
_______ (1983), *Logic, Semantics, Metamathematics*, 2nd. edn., John Corcoran (ed.), translations by J. H. Woodger, Hackett Publishing Company (Indianapolis).
_______ (1986) *Collected Works*, vols. 1-4, (S. R. Givant and R. N. McKenzie, eds.), Birkhäuser (Basel).
_______ (1986a), What are logical notions?, *History and Philosophy of Logic 7*, 143-154.
A. Tarski and S. Givant (1999), Tarski's system of geometry, *Bull. Symbolic Logic 5*, 175-214.

A. Tarski and R. L. Vaught (1957), Arithmetical extensions of relational systems, *Compositio Mathematica 13,* 81-102.

J. van Heijenoort (ed.) (1967), *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931*, Harvard University Press (Cambridge). (Third printing, 1976.)

______________ (1967a), Logic as calculus and logic as language, *Boston Studies in the Philosophy of Science 3*, Reidel Pub. Co., 440-446; reprinted in J. van Heijenoort (1985), 11-16.

______________ (1976) Set-theoretic semantics, in *Logic Colloquium '76* (R. O. Gandy and M. Hyland, eds.), North-Holland (Amsterdam), 183-190; reprinted in van Heijenoort (1985), 43-53.

______________ (1985) *Selected Essays*, Bibliopolis (Naples).

R. L. Vaught (1986), Tarski's work in model theory, *J. Symbolic Logic 51,* 869-882.

___________ (1974), Model theory before 1945, in Henkin et al. (1974), 153-172.

H. Weyl (1910), Über die Definitionen der mathematischen Grundbegriffe, *Mathematisch-naturwissenschaftliche Blätter 7*, 93-95 and 109-113; reprinted in H. Weyl (1968), *Gesammelte Abhandlungen,* Vol. I, 298-304.

J. Wolenski (1989), *Logic and Philosophy in the Lwow-Warsaw School,* Kluwer Academic Publishers (Dordrecht).