Logic, Epistemology, and the Unity of Science 45

Jan Woleński

Semantics and Truth

Springer

Logic, Epistemology, and the Unity of Science

Volume 45

Logic, Epistemology, and the Unity of Science aims to reconsider the question of the unity of science in light of recent developments in logic. At present, no single logical, semantical or methodological framework dominates the philosophy of science. However, the editors of this series believe that formal techniques like, for example, independence friendly logic, dialogical logics, multimodal logics, game theoretic semantics and linear logics, have the potential to cast new light on basic issues in the discussion of the unity of science.

This series provides a venue where philosophers and logicians can apply specific technical insights to fundamental philosophical problems. While the series is open to a wide variety of perspectives, including the study and analysis of argumentation and the critical discussion of the relationship between logic and the philosophy of science, the aim is to provide an integrated picture of the scientific enterprise in all its diversity.

For inquiries and submissions of proposals, authors can contact Christi Lue at christi.lue@springer.com

More information about this series at http://www.springer.com/series/6936

Jan Woleński

Semantics and Truth

Springer

Jan Woleński
Jagiellonian University (prof. emeritus)
Kraków, Poland

University of Information, Technology
and Management
Rzeszow, Poland

ISSN 2214-9775 ISSN 2214-9783 (electronic)
Logic, Epistemology, and the Unity of Science
ISBN 978-3-030-24535-1 ISBN 978-3-030-24536-8 (eBook)
https://doi.org/10.1007/978-3-030-24536-8

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To my Family

Preface

This book has a long history. I began to work on it in 2001, on the occasion of investigations supported by the Committee of Scientific Research in Poland at the end of the last century, realizing a grant. I continued the research on the concept of truth during my stay at the Netherlands Institute of Advanced Study in Waassenaar from September 2003 to June 2004. I would like to express my deep thanks to these institutions. Unfortunately, teaching and other professional duties caused me to interrupt writing this book, although I published several papers on truth (historical as well as substantial) in the interim. I summarized many of these results in Woleński 2005 (a book on epistemology and its problem—in Polish). I use a substantial portion of the material already published in the present monograph as well as the mentioned papers on truth; details will be provided at appropriate places, but I wish to give a special mention to my monograph Woleński 1989 on the Lvov–Warsaw School which provides a description of the general philosophical environment relevant for the topic of the present book.

Technical remarks. This book has no endnotes or footnotes. I belong to that group of readers who dislike the latter and can barely tolerate the former. Instead, I have introduced digressions, indicated by (**DG**); they end with the sign ►. Every chapter has its own numbered set of digressions with indications of the type (**DGn**), where the digit refers to the number of a given digression. References to digressions are indicated by the sequence of the type **DGnX**, where the indication to following the digit refers to the number of the chapter in which the given digression occurs; if the reference concerns digression in the same chapter, the sequence **DGn** is employed. For example, the sequence **DG1VII** refers to the first digression in Chap. 7 and the sequence **DG3**—to the third digression in a given chapter. The same convention applies to references made to numbered formulas and definitions, where they occur. References to sections (§), formulas, digressions, and definitions within the same chapter omit its number. I freely use devices to indicate distinctive phrases (mainly formulas and lists of questions), namely, numerals, letters, or special signs, like # or (*). In general, such references apply to particular paragraphs, but I hope that specific contexts preclude possible misunderstanding. Here is an example of a digression:

(**DG1**) One may wonder how I use the word 'theory' with respect to a set of philosophical statements. Philosophy does not offer collections of sentences as being either logical or mathematical theories—that is, sets of sentences closed by the consequence operation or as empirical theories, namely, collections of hypotheses formulated in order to explain or predict some empirical data. As I shall show in Chap. 1, aletheiology (this word is derived from the German neologism *Alethiologie*, introduced by Johannes Lambert in the eighteenth century; 'aletheiology' can be regarded as a substitute for 'philosophy of truth') has to fulfill some tasks stemming from its history. Such an enterprise always leads to a definite class of statements that answer traditional questions. Such answers are traditionally called truth-theories or theories of truth, and there is no reason to abandon this terminology. In general, philosophical theories are bodies (classes, complexes, sets, etc.) of interconnected statements, which are subjected to philosophical and metaphilosophical constraints—for example, that we work *via* claims, like that every philosophical problem is legitimate, provided that it can be naturalized.►

Bibliographical references consist of the author's name, the publication date, and/or page-number(s) (note that the sequence M, N 2000 refers to names of co-authors or co-editors of a joint piece published in 2000)—except for classical sources up to and including Kant. The latter are quoted or mentioned in full (first time), with information about English translation, if any, or by abbreviated (in further cases) title. If a translator's name is not given, translation is mine. These sources are not included in the bibliography at the end of the book. This way of treating the classical sources is motivated by my feeling that brief references to them such as Kant 1787 are odd. Moreover, the exact dates of some of the older sources are unknown. Pre-Socratics are mainly quoted after H. Diels, *Fragmente der Vorsokratiker*, 3 vls., 17th ed. Berlin: Weidmannsche Verlagsbuchhandlung 1954 (I use the standard notation: Diels I 4B 35 refers to fragment 35 in the section B of chapter 4 of volume 1) or G. S. Kirk, J. E. Raven, M. Schofield, *The Presocratic Philosophers*, 2nd ed., Cambridge: Cambridge University Press 1983 (references: Kirk, Raven, Schofield plus page-number); otherwise, a special information is provided. In other cases (until the nineteenth century), I usually mention the title and chapter (section, etc.). For historical reasons, references (to writings included in the bibliography) are almost always to originals and first editions (very few exceptions are justified by the lack of historical relevance or proximity of particular editions). Consequently, the titles of the books and papers listed in the bibliography at the end are always given in the language of the original. If an English translation or a later edition is also mentioned in the bibliography, page-references are to it. Names of translators are provided only in the case of quoting from a given book (paper). If several writings are cited together, they are chronologically ordered.

One more remark about bibliographical data is in order. There are thousands (that is no exaggeration) of writings about truth and problems related to this concept. Also Tarski's truth-theory, which is the main focus of my considerations, was presented, commented on, and criticized so many times that it is difficult to assess their number. So I had to make a selection of quoted books and papers, but I decided to include into the Bibliography rather a long list of my own contributions related to the concept of truth. The reason is not that I try to promote myself, but to give a credit to the already published material used in this book. In general, I frequently omit references in the case of marginal questions, illustrative examples, or commonly known facts from the history of logic and philosophy. Clearly, my decisions are to some (or even a great) extent subjective. I apologize in advance for all bibliographical inaccuracies.

One omission should be especially mentioned. I resigned from quoting many Polish books and papers. On the other hand, I am greatly indebted to many of my Polish friends and colleagues for stimulating discussions and/or benefit that stemmed from reading their writings. The list of these persons is included in alphabetical order: Anna Brożek, Wojciech Buszkowski, Bogdan Chwedeńczuk, Roman Duda, Katarzyna Gan–Krzywoszyńska, Adam Grobler, Michał Heller, Jacek J. Jadacki, Elżbieta Kałuszyńska, Anna Kanik, Katarzyna Kijania–Placek, Sebastian Kołodziejczyk, Stanisław Krajewski, Piotr Łukowski, Grzegorz Malinowski, Witold Marciszewski, Wiktor Marek, Roman Murawski (particularly for his consultations about formal matters), Jan Mycielski, Adam Nowaczyk, Adam Olszewski, Leszek Pacholski, Jacek Paśniczek, Tomasz Placek, Jerzy Pogonowski, Michael Schudrich, Andrew Schumann, Marcin Selinger, Stanisław J. Surma, Jerzy Szymura, Marek Tokarz, Kazimierz Trzęsicki, Urszula Wybraniec–Skardowska, Ryszard Wójcicki, Andrzej Wroński, and Jan Zygmunt. Although I credited my debts in quotations, some persons from the abroad deserve to be especially mentioned for their remarks and discussions with them, namely, Joseph Agassi, Evandro Agazzi, David Armstrong, Nuel Belnap, Arianna Betti, Natan Berber, Jean-Yves Béziau, Johannes Brandl, Maria Luisa Dalla Chiara, Franco Conglione, John Corcoran, Marian David, Michael Devitt, Pascal Engel, Susan Haack, Hartry Field, Keith Fine, Juliet Floyd, Dagfinn Føllesdal, Paul Horwich, David Kashtan, Wolfang Künne, Eckerhardt Köhler, Saul Kripke, Kevin Mulligan, Ilkka Niiniluoto, David Pearce, Volker Peckhaus, Roberto Poli, Carl Posy, Michael Potter, Gabriel Sandu, Denis Savieliev, Dana Scott, Krister Segerberg, Valentin Shehtman, Gila Sher, Peter Simons (we co-authored together the paper important for this book), Barry Smith, Göran Sundholm, Matti Sintonen, Jan Tarski, Christian Thiel, Max Urchs, Jean-Yves Beziau, Jan von Plato, and Paul Weingartner. I also cannot omit to mention colleagues and friends who passed away, in particular, Józef M. Bocheński, Donald Davidson, Burton Dreben, Solomon Feferman, Andrzej Grzegorczyk, Rudolf Haller, Jaakko Hintikka, Henryk Hiż, Jerzy Kalinowski, Stig Kanger, Alexander Karpenko, Czesław Lejewski, Jerzy Łoś, Leszek Nowak, Jerzy Pelc, Jerzy Perzanowski, Ingmar Pörn, Marian Przełęcki, Hilary Putnam, Willard v. O. Quine, Barbara Stanosz, Roman Suszko, Klemens Szaniawski, Aleksander Szulc, Georg Henrik von Wright, and Józef Życiński. All lists included above are surely incomplete and please forgive me for omissions.

Springer Verlag agreed to publish the present book in the series “Trends in Logic” (Heinrich Wansing, the editor) at first, and finally—“Logic, Epistemology, and the Philosophy of Science” (Shahid Rahman, the editor). I am very indebted to the publisher, and both mentioned editors for patience allowing me to complete this

book. Adam Tuboly kindly provided English translations of Otto Neurath's letters to Rudolf Carnap quoted in Chap. 9, and permitted me to quote them. Romina Padro sends me Kripke's forthcoming paper (see Bibliography). Last but not least, I am very indebted to Arthur Szylewicz for thorough job of emending my English, to Jacqueline Duong Nguyen for processing those changes electronically, and to the anonymous referee for his/her valuable remarks. Although I did not follow all suggestions of the referee, the final version of the text is certainly better than the earlier one.

All non-English single words or nominal phrases are printed in italics, and usually occur without quotation marks—except for those that occur in cited passages; the same applies to Greek or Latin philosophical maxims, for example, the famous sentence *ens et bonum convertuntur*. Quotations are normally inserted as separate fragments printed in smaller letters (non-English fragments are printed in italic in such cases), others occur in double quotes ("…"); such quotes are also used to mark a metaphorical meaning of a given phrase. Single quotes ('…') indicate that an expression is mentioned, but not used (see **DGIII3** for an explanation of this distinction). In order to avoid using quotes too frequently, I adopt the standard convention that such phrases as 'the expression …', 'the letter …', 'the variable …', 'the formula …', etc. indicate that their completion stands in the material mode, that is, are mentioned, not used. Thus, the phrase 'the variable *x*' abbreviates 'the variable '*x*'. However, this convention (following the style employed by Polish logicians) applies only to the fully symbolic contexts. If a phrase contains solely words, or words and symbols, we write, for example, 'the sentence 'snow is white', but not 'the sentence snow is white' and 'the formula '*x* is white', but not 'the formula *x* is white'. All citations preserve the original, also its way of employing quotation marks, with the exception that double-spaced print, sometimes occurring in older German writings, is replaced by italics (for instance, 'N a m e' by '*Name*'). The African, Chinese, Greek, Hebrew, and Sanskrit words occur in simplified Latin transcriptions. I am fully aware that technical rules prescribed in this book are conventional, and that their use sometimes looks artificial, but I hope these circumstances do not lead to misunderstandings.

Kraków, Poland
Rzeszow, Poland

Jan Woleński

Reference

Woleński, J. (1989). *Logic and Philosophy in the Lvov–Warsaw School*. Dordrecht: Kluwer.

Introduction

Abstract Presenting Tarski's semantic theory of truth (**STT**) as a formal logical construction and as a philosophical theory is the task of the book. Although **STT** as a formal theory is commonly recognized, its philosophical significance is debated. Several opinions of logicians and philosophers are quoted in order to show the state of art of the discussions around **STT**. My own attitude considers **STT** as philosophically important. Moreover, I explain my analytical methodology consisting of so-called interpretative consequences.

This book offers a systematic exposition of the semantic theory of truth (**STT** henceforth) in frameworks of semantics and logic. This theory, formulated by Alfred Tarski in 1930s, has two separate, though closely interconnected, aspects. First, **STT** is a formal logical (or even mathematical) theory and functions as the central conceptual foundation of model theory, next to proof theory and recursion theory, of the most important branches of modern mathematical logic. Second, **STT** is also a significant philosophical doctrine (see Woleński 1999a), which tries to elaborate the notion of truth as investigated by philosophers from antiquity to contemporary times. The assessment of **STT** as a mathematical theory on the one hand, and as a philosophical doctrine on the other, is however different to some extent. Consider the following prophecy (Hodges 1985–1986, p. 135):

> But before you dismiss him as a mere theorem prover, you should ask yourself what your grandsons and granddaughters are likely to study when they settle down to their 'Logic for computing class' at 9.30 after school assembly. Will it be syllogisms? Just possibly it could be the difference between saturated objects and unsaturated concepts, though I doubt it. I put my money on Tarski's definition of truth for formalized languages. It has already reached the universal textbooks of logic programming, and another ten years should see it safely into the sixth forms. This is a measure of how far Tarski has influenced the whole framework of logic

Clearly, in the quoted fragment, Hodges talks about **STT** as a well-established mathematical theory. Independently of whether Hodges' prophecy is right, or perhaps too optimistic with respect to the education of our grandsons and granddaughters, Tarski's truth-definition is permanently in vogue among mathematical logicians and specialists in the foundations of mathematics, and almost nobody denies its importance as an idea within mathematical logic. If the reader wonders why I say "almost nobody", I would like to recall what Turing said once about **STT**, namely, that "Triviality can go no further" (see Wang 1986, p. 144). Turing's words elicited the following view from Hao Wang (p. 144):

> There is a great difference of opinion on the importance of [Tarski's] contribution to this area [that is, the theory of truth—J. W.].

It is not quite clear whether this evaluation concerns the formal aspect of **STT** or its philosophical content or even both. Nevertheless, it is fair to say that the importance of Tarski's work as a mathematical enterprise is much closer to Hodges' view than to Turing's and Wang's opinion.

That Tarski himself considered **STT** as a philosophical doctrine can be clearly documented by two passages taken from his main work (Tarski 1933, p. 152, pp. 266–267; the first opens the book, the second almost closes it):

> The present article is almost wholly devoted to a single problem—*the definition of truth*. Its task is to construct—with reference to a given language—*a materially adequate and formally correct definition of the term 'true sentence'*. This problem [...] belongs to the classical questions of philosophy [...].
> [...] in its essential parts the present work deviates from the mainstream of methodological study [that is, metalogical or metamathematical; the scope of the methodological study should be seen here in a wider sense than in the Hilbert school, that is, as not restricted to finitary proof theory—JW]. Its central problem—the construction of the definition of true sentence and establishing the scientific foundations of the theory of truth—belongs to the theory of knowledge and forms one of the chief problems of philosophy. I therefore hope that this work will interest the student of the theory of knowledge [in the Polish original "zainteresują się przede wszystkim teoretycy poznania", which literally means "will interest above all epistemologists"—JW] that he will be able to analyse the results contained in it critically and to judge their value for further research in this field, without allowing himself to be discouraged by the apparatus of concepts and methods used here, which in places have been difficult and have not been used in the field in which he works."

However, **STT** as a philosophical doctrine is far more complex and there certainly is—to repeat Wang's evaluation—for the most part proper in this context, "a great difference of opinion on the importance of [Tarski's] contribution." To start with positive responses, Tarski's ideas became immediately welcomed by philosophers using logical tools in philosophical investigations ('logical philosophers' is a label that has recently gained popularity). Alfred Ayer wrote (Ayer 1967, p. 116):

> Philosophically the highlight of the Congress [in Paris in 1935—J. W.] was the presentation by Tarski of a paper which summarized his theory of truth.

Three important contemporary philosophers, namely, Kazimierz Ajdukiewicz, Rudolf Carnap, and Karl Popper radically changed or at least modified their earlier views under Tarski's direct influence. Ajdukiewicz abandoned radical conventionalism, which was, among other things, a theory of language and meaning (Ajdukiewicz 1964, p. 315):

> The objection […] communicated to me by Tarski in a conversation […] seems to show that the concept of meaning is not definable in purely syntactical terms without the use of semantic terms in the narrower sense.

Carnap made a similar point (Carnap 1942, p. X):

> Tarski, both through his book, and in conversation, first called my attention to the fact that the formal method of syntax and semantics must be supplemented by semantic concepts, showing at the same time that these concepts can be defined by means not less exact than those of syntax. Thus the present book owes very much to Tarski, more indeed than to any other single influence.

Briefly, Carnap passed, under Tarski's influence, from philosophy as logical syntax to philosophy as exact semantic analysis. It is no exaggeration to say that Tarski made an essential contribution to the semantic revolution in philosophy (see Woleński 1999b and Chap. 6).

Finally, Popper recalls (Popper 1972, p. 322; see also Hazohen 2000, passim):

> […] I met Tarski in July 1934 in Prague. It was early in 1935 that I met him again in Vienna in Karl Menger's Colloquium […] It was in those days that I asked Tarski to explain me his theory of truth, and he did so in a lecture of perhaps twenty minutes on a bench (un unforgotten bench) in the *Volksgarten* in Vienna. He also allowed me to see the sequence of proofs sheets of the German translation of his great paper on the concept of truth, which was than just sent to him from […] *Studia Philosophica*. 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 learn so much from anybody else.

How did Tarski's ideas influence Popper? Generally speaking, Popper abandoned his earlier doubts about the concept of truth and adopted realism in his approach to science. In particular, he came to the conclusion that **STT** rehabilitated the idea that truth consists in conformity of propositions to objective reality.

These three examples of the acceptance of Tarski's ideas together along with Ayer's general assessment are perhaps the most spectacular traces of Tarski's influence on philosophy. However, the philosophical role of **STT** is by no means limited to these specific works. Almost every book (introductory or advanced) in semantics, philosophy of language, or the history of analytic philosophy gives a summary of or, at least, mentions it. Similarly, almost every discussion of how to define meaning, semantic realism, or scientific realism employs Tarski's results, or at least alludes to them. Several important views in contemporary philosophy make use **STT**, for example, Donald Davidson's theory of meaning as based on truth-conditions (see Chap. 9, Sect. 9.4) or various semantic theories of induction (Carnap and his followers). Tarski's theory was more or less modified, like in Kripke 1975 or Gupta, Belnap 1993, or replaced by other constructions, as in Hintikka 1996. Since both modifications and replacements refer to **STT** as the solid starting point, it can be generally said that Tarski's ideas attracted many leading philosophers, contributed to the semantic revolution, gave the rise to several modifications and constructions regarded as alternatives to the semantic theory of truth, stimulated investigations on a variety of philosophical problems and, last but not least, found a lasting place in textbooks, monographs and anthologies. It is no exaggeration that every post-Tarskian theory of truth (at least in analytic philosophy), even if critical to some extent, is propter-Tarskian. Saul Kripke expressed this dependence by saying (Kripke 1975, p. 97) that the ghost of the Tarski hierarchy (of languages; see Chaps. 7–8) "is still with us." (see Kripke 2019a, for a more sophisticated, than in Kripke 1975, treatment of the issue of language-hierarchies).

The above focuses on the positive influence of Tarski's ideas as something accepted, or at least stimulating, in philosophical investigations. However, **STT** is also strongly criticized. Of course, it is not surprising that most non-analytic philosophers, of the post-modernist camp, for example, simply ignore this theory, or even regard it as a typical degeneration of the logical mind. I will not comment on such criticisms, although I would like to explain why. A discussion between philosophers belonging to various philosophical camps is a delicate matter. The main problem is that metaphilosophical options contribute substantially to resolving issues. Thus if someone says as Martin Heidegger does, that truth is entirely outside logic or semantics and must be located in philosophical anthropology, there is very little chance of a fruitful discussion between such a philosopher and one who believes philosophy to be based on logical analysis. As a dedicated logical philosopher, I do not say that other philosophies are wrong and have no value. I only indicate that, except to register fundamental metaphilosophical contrasts and their effects, I do not have very much to discuss with non-analytical or post-analytical philosophers; their attitude will be similar, of course. A consequence of this view, which I regard as rational, leads to the claim that I will focus on criticisms of **STT** that arose inside the analytical camp or its vicinities. Since various arguments advanced for by analytic philosophers against will be discussed in many places of this book (particularly in Chap. 9), at this point I note only a handful of examples. Max Black (see Black 1948) tried to show that **STT**, although correct from a purely logical point of view, is neutral in fact with respect to old philosophical controversies about the concept of truth. Perhaps the most radical criticism of **STT** is that of Hilary Putnam (see Putnam 1975a, Putnam 1983, Putnam 1985–1986). He argues that **STT** theory, although proper for mathematical logic, is incorrect as a philosophical proposal and deceives philosophers. Yet objections against **STT** strongly suggest that Tarski was effectively achieving his goal to interest philosophers in his ideas. When we browse the Internet, we find virtually tens of thousands to Tarski and his theory truth. Admittedly, this is considerably fewer than when we search 'Heidegger and truth' (almost sixty thousands), but this last topic is much broader and accessible to everyone with philosophical ambitions, whereas discussing **STT** requires some specialized knowledge and logical competence.

In spite of the fact that **STT** is located at the heart of (analytical) philosophy, there is as yet no comprehensive systematic stud on it. Of course, there are various treatments. Some are long, other shorter, some are more technical, other less technical, some are simplified other advanced, but none, at least as far as I know, try to deal with all or the main philosophical problems related to **STT**. The present monograph tries to fill this gap. It is intended as a multifaceted philosophical study of **STT**. I previously noted that the formal mathematical aspects of **STT** and its philosophical features are interconnected. However, their mutual interplay is not symmetrical. If one sketches or even fully elaborates **STT** as a part of model theory in mathematical logic, one does not need to allude to the philosophical content of the theory. Such a practice has become the norm in contemporary textbooks and monographs on logic and model theory (see, for example, Enderton 1972, Chang, Keisler 1973, Doets 1996, Manzano 1999, Hinman 2005). This situation is not surprising, as the content of mathematical theories is usually independent (and it should be) of their philosophical background.

However, the reverse, that is, the direction from philosophy to logic, is different, according to metaphilosophical principles I share. Formal (logical) philosophical analysis cannot be independent of the technical results of logic. Let me use an analogy to explain the point. We can debate about determinism, indeterminism, and related topics without any appeal to physics. Nevertheless, it seems pointless to discuss these issues while ignoring quantum mechanics and the physical theory of chaotic phenomena. Similarly, it is perfectly possible to discuss the concept of truth without any appeal to logic, metamathematics, and formal semantics. This analogy goes further. Suppose that we want to speak about the philosophical consequences of Heisenberg's uncertainty principle. In particular, we want to investigate whether the formula $\Delta p_1 \bullet \Delta p_2 \geq h$ (the product of indeterminacies of momentum and position of an elementary particle is greater than the Planck constant; this formulation is simplified with respect to h) entails indeterministic consequences, or not.

If we take the word 'entails' in its strict logical sense, a discussion pertaining to deterministic or indeterministic consequences of Heisenberg's principle is simply not possible. The reason is that the terms 'determinism' and 'indeterminism' (or related adjectives) do not occur in the formulation of the principle. In order to derive an ontological statement about the nature of the world, we need to embed the uncertainty principle into the philosophical vocabulary. Heisenberg himself did this by using the frequently held view of determinism which claims that the future can be predicted if we have an exact knowledge of the present state of reality. Since the uncertainty principle essentially precludes an exact knowledge of the present state of reality, deterministic predictions are impossible. Ergo, indeterminism is correct. However, other philosophical embeddings are also possible, for instance, weakened determinism and indeterminism. If we see determinism as consistent with statistical or probabilistic predictions, the relation between the uncertainty principle and the deterministic structure of reality becomes more complicated than under Heisenberg's view. Hence, we can conclude that the physical sense of the uncertainty principle is completely independent of the philosophical embeddings imposed on it. Thus, the philosophical consequences of the Heisenberg principle do not derive directly from it, but from its reformulations, relative to adopted philosophical interpretations. I qualify such conclusions derived from scientific results as interpretative consequences. In particular, indeterminism is an interpretative consequence of the Heisenberg principle when the mentioned interpretation of determinism is accepted, that is, when the uncertainty principle is seen modulo the idea that the future can accurately be predicted from information about the past. In order to obtain interpretative consequences of scientific statements, the Heisenberg principle, for example, one should embed the latter into a philosophical language. Note that such embeddings should not be considered as exact translations.

The idea of interpretative consequences accords very well with a vision of philosophy in its (chosen) analytic setting in particular. I agree in principle with the following view (Waismann 1956, p. 1):

> […] philosophy, as it is practised today, is very unlike today, is very unlike science; and this in three respects: in philosophy there are no proofs; there are not theorems; and there are no questions which can be decided, Yes or Not. In saying that there are no proofs I do not mean to say that there are no arguments. Arguments certainly there are, and first-rate philosophers are recognized by the originality of their arguments; only these do not work in the sort of way they do in mathematics or in the sciences.

Observe that there is a contradiction between Waismann's view and the idea of interpretative consequences because the latter does not preclude that philosophical problems have the answers: Yes or No. But the point is that interpretative consequences do not work as scientific arguments. On the other hand, the suggested method of analysis *via* philosophical embeddings of various scientific—in particular, mathematical and physical results—and deriving interpretative consequences from them shows how the philosophical arguments proceed and provide means for their evaluation. For example, I am inclined to regard the arguments for teleology derived from Aristotle's physics as obsolete and wrong, whereas I see criticism of these arguments based on the theory of evolution or the theory of chaotic phenomena as sound. However, I have no tools to demonstrate that relevant philosophical embeddings are absolutely incorrect, because, for example, no empirical investigation can justify the view that Aristotle's theory of substance is wrong. Thus, a Thomistic philosopher can always say that he or she intuitively sees substances as composed of form and matter and there is no way to convince them that this idea is wrong. All we can do is argue that Aristotle's vision of substance is at odds with physics, but the Aristotelians can always defend their position by pointing out that philosophy is more fundamental than natural science. The gap between various (meta)philosophical camps is indeed very wide (or deep, if you prefer this way of speaking about philosophical issues).

I will consider **STT** not only as a piece of philosophy (it is out of the question) but also as good philosophical theory (it is problematic). I will argue, as Tarski himself did, that **STT** not only remains inside the definite Aristotelian tradition but also illuminates it in a very interesting way. My argumentation will proceed *via* the interpretative consequences derived from the philosophical embeddings imposed on the logical machinery employed in **STT**. Hence, this monograph takes **STT** seriously as a formal theory. One can now ask for the source of philosophical embeddings (interpretations) that generate interpretative consequences. Although it is not an easy process, the best place to look for insights in this respect is the history of philosophy. We need to look to history for the investigation of any truth-theory, because the problem of truth certainly has been one of the philosophical invariants, since Aristotle at least. In the philosophy of truth, as in other branches of philosophy, the basic collection of problems originated from the ancient Greeks. Generations of philosophers have worked on the concept of truth, often producing entirely new insights. As is customary in philosophy, some questions disappear and some reappear while new ones emerge. One can ask why logic is an important source of philosophical ideas. My answer follows Stanisław Leśniewski's view (Henry Hiż's personal communication) that logic is a formal exposition of intuition.

The context described above determines the structure of the present book, which has substantive as well as historical ambitions. I begin with three chapters on the history of the concept of truth from antiquity up to the nineteenth and twentieth centuries. It is astonishing that truth, one of the most important concepts in all philosophy, still awaits a full historical exposition (see Enders 1999, Szaif 2006 and Pritzl 2010 for a partial realization of this task). The temptation to redress this imbalance and write a complete history of the truth-concept was great. However, I decided to limit the historical side of my study to an investigation of the classical or correspondence theory of truth, although, as I will show later, we need to distinguish between the classical theory of truth and the correspondence theory of truth. Other theories are mentioned only in passing. A special section (in Chap. 3) is devoted to Polish works on truth, because Tarski grew up in a specific philosophical environment determined by the ideas of Kazimierz Twardowski and his followers (the Lvov–Warsaw School) and because this intellectual climate essentially influenced the content of **STT**. Two issues arose in connection with the subject matter of the historical chapters. First, although it is true that in philosophy (at least) everything can be compared with something else—and therefore we could compare **STT** with the pragmatic theory of truth or the consensus theory—such a procedure would be pointless, because the related sets of ideas are fundamentally different. Second, I decided to include a review of many historical points in order to show that **STT** belongs to the trajectory of arguments which regard truth as consisting in saying that something is so and so and something is just such and such. I hope that the historical part of my study, in spite of its shortcomings and incompleteness, possesses some autonomous value as an introduction to a more ambitious history of aletheiology. Anyway, if history is considered as the teacher, this role of it is as important in philosophy as in elsewhere. Chapter 4 outlines the tasks that form the basis for any philosophical theory of truth. This fragment is quite straightforward, as I first wanted to focus on some basic concepts for explaining some preliminary issues, and to introduce the most important currents of thinking within the past and present philosophy of truth.

Chapter 5 presents the logical basis of my further analysis (some logical problems are also considered in Chap. 4, but in a semi-formal manner). I touch on various logical and metalogical topics in order to provide formal tools for a more advanced analysis of **STT**. I decided to present the rudiments of logic and metalogic for three reasons. First, I want to make this book self-contained. Second, I wanted to set uniform terminological usages employed in further parts of the book. Third, formal concepts and results provide the instruments to facilitate a discussion of (some) philosophical aspects and the uses of **STT**. Matters of semantics are discussed in Chap. 6. Two next chapters contain informal (Chap. 7) and formal (Chap. 8) presentation of **STT**. In particular, an explicit picture of the relation between syntax and semantics is an outcome of limitative theorems presented and discussed in Chap. 8. That semantics is not reducible to syntax, I consider perhaps as the most important moral coming from the analysis of **STT**. Chapter 9 discusses some interpretative, comparative, and philosophical issues related to the semantic theory of truth. A more detailed survey of the content of the last chapter is provided in introduction to it. The book ends with a short conclusion concerning the status of **STT** as a piece of philosophical analysis.

I intend to follow Tarski's way of formulating **STT** rather closely. In particular, I propose to take his arguments seriously and I defend most of his views. On the other hand, Tarski is not sacrosanct and some of his views must (or should be) be corrected. I mention two departures from the original version of **STT**. The first concerns the assumed formalism (I will repeat these remarks in other places of the book). Tarski formulated his truth-theory for a version of the simple theory of types. Contemporary textbooks and monographs employ the first-order logic and its metalogic. Since this change agrees with Tarski's suggestions implicit in his later works, it can be regarded as of a secondary importance. A more essential departure concerns the philosophical content and consequences of **STT**. Tarski was very careful in expressing his philosophical views (see Mostowski 1967, Suppes 1988) and usually abstained from articulating them, particularly, in his writings. On the other hand, he was more ready to speak about philosophical issues in oral discussions, but not very much information preserved (see Feferman, Feferman 2004 for perhaps the most extensive documentation). Tarski's attitude toward philosophical declarations does not allow to reconstruct his views about many interesting philosophical issues provoked by **STT**. I decided to say much more about these questions, because I believe that the philosophical content of Tarski's theory is more comprehensive than he admitted.

References

Black, M. (1948). The Semantic Definition of Truth. Analysis, 8, 49–63; repr. (with additions), in Black (1949), 89–107, 251.

Black, M. (1949). *Language and Philosophy*. Ithaca: Cornell University Press

Cauman, L. S., Levi, I., Parsons, Ch., Schwartz, R. (Eds.) (1983). *How Many Questions? Essays in Honour of Sidney Morgenbesser*. Indianapolis: Hackett Publishing Company.

Doets, K. (1996). *Basic Model Theory*. Stanford: Center for the Study of Language and Information.

Edwards, P. (Ed.) (1967). *The Encyclopedia of Philosophy*, v. 1–8. New York: Macmillan.

Enderton, H. B. (1972). *A Mathematical Introduction to Logic*. New York: Academic Press.

Enders, M. (1999). *Wahrheit und Notwendigkeit. Die Theorie der Wahrheit bei Anselm of Canterbury im Gesamtzusammenhang seines Denkens und unter besonderer Berücksichtigung seiner antiken Quellen (Aristoteles, Cicero, Augustinus, Boethius)*. Leiden: Brill.

Enders, M., Szaif, J. (2006). *Die Geschichte des philosophischen Begriffs der Wahrheit*. Berlin de Gruyter.
Feferman, A., Feferman, S. (2004). *Alfred Tarski. Life and Logic*. Cambridge: Cambridge University Press.
Gupta, A., Belnap, N. (1993). *The Revision Theory of Truth*. Cambridge, Mass.: The MIT Press.
Hazohen, M. H. (2000). *Karl Popper. The Formative Years 1902–1945*. Politics and Philosophy in Interwar Vienna. Cambridge University Press: Cambridge.
Hinman, P. G. (2005). *Fundamentals of Mathematical Logic*. Wellesday, Mass.: A. K. Peters.
Hintikka, J. (1996). *The Principles of Mathematics Revisited*. Cambridge: Cambridge University Press.
Kripke, S. (1975). An outline of a theory of truth. The Journal of Philosophy, 72, 690–716; repr. in Kripke (2011), 75–98.
Kripke, S. (2011). *Philosophical troubles. Collected Papers*, v. 1, New York: Oxford University Press.
Kripke, S. (2019). *Philosophical Troubles. Collected Papers*, v. 2. New York: Oxford University Press (forthcoming).
Kripke, S. (2019a). Ungroundedness in Tarskian languages. In Kripke (2019).
Lewis, D. (Ed.) (1956). *British Philosophy*, *III*. London: George Allen and Unwin.
Manzano, M. (1999). *Model Theory*. Oxford: Clarendon Press.
Mostowski, A. (1967). Alfred Tarski. In Edwards (1967), v. 8, 77–81.
Peregrin, J. (Ed.) (1999). *The Nature of Truth (If Any)*. Dordrecht: Kluwer.
Pritzl, K. (Ed.) (2010). *Truth. Studies in a Robust Presence*. Washington, D. C.: The Catholic University of America Press.
Putnam, H. (1975). *Mind, Language and Reality. Philosophical Papers*, v. 2. Cambridge: Camrbidge University Press.
Putnam, H. (1975a). Do true assertions correspond to reality? In Putnam (1975), 70–84.
Putnam, H. (1983). On truth. In Cauman, Levi, Parsons, Schwartz 1(983), 35–56; repr. in Putnam, H. (1994), 315–329.
Putnam, H. (1985–1986). On comparison of something with something else. *New Literary History*, 17, 61–79; repr. in Putnam (1994), 330–350.
Putnam, H. (1994) *World and Life*. Cambridge, Mass: Harvard University Press.
Suppes, P. (1988). Philosophical implications of Tarski's work. *The Journal of Symbolic Logic*, 53, 80–91.
Waismann, F. (1956). How I see philosophy? In Lewis (1956), 447–490; repr., in Waismann (1968), 1–38.
Waismann, F. (1968). *How I see Phiolosophy*. London: Macmillan.
Woleński, J. (1989). *Logic and Philosophy in the Lvov–Warsaw School*. Dordrecht: Kluwer.
Woleński, J. (1999a). Semantic conception of truth as a philosophical theory. In Peregrin, (1999), 51–66.
Woleński, J. (1999b). The semantic revolution—Rudolf Carnap, Kurt Gödel, Alfred Tarski. In Woleński, Köhler (1998), 1–15.
Woleński, J. (2005). *Epistemologia. Poznanie, prawda, wiedza, realizm* (Epistemology. Cognition, Truth, Knowledge and Realism). Warszawa: Wydawnictwo Naukowe PWN.
Woleński, J., Köhler, E. (Eds.) (1999). *Alfred Tarski and the Vienna Circle. Austro-Polish Connection in Logical Empiricism*. Dordrecht: Kluwer.

Contents

Chapter 1
Truth in Ancient Philosophy

Abstract This chapter opens the historical part of the book. I focus on Greek philosophy (archaic poets and thinkers, Pre-Socratics, Plato, Aristotle, post-Aristotelian, particularly the Stoics), some facts from other philosophical cultures (Chinese, Indian, African, Egyptian, Hebrew) are mentioned. The historical report in this and next chapters (see also Woleński 1994a) tries to show that the understanding of truth as saying how things are, is present in all cases taken into account.

1.1 Archaic Greece

(**DG1**) In later digressions in this section I go beyond Greek ancient thought. In particular, I consider ideas present in old German language, as well as in Indian religions, Judaism, Islam, and Chinese thought. The reason is that they are not associated with explicitly articulated (or fully separated from religion) philosophy and thereby more similar to the archaic Greek thinking than to the professional philosophy of later Greek thinkers, including the Pre-Socratics. On the other hand, data from various more or less ancient languages, seem to confirm my thesis that **STT** is deeply rooted in a very colloquial understanding of the predicate 'is true' (see Woleński 2004f.).►

The origin of philosophy and science in Greece is still regarded as a singular event, even a kind of miracle. Hence, we encounter various explanations of this fact. There are three principal theories at issue. According to John Burnet (see Burnet 1892), Greek philosophy appeared as a result of a transition (Burnet sees it as somehow mysterious) from archaic ('archaic' means here 'pre-philosophical') myths and religion to secular thought. Another theory was advanced by Francis M. Conford (see Conford 1952) who maintained that the process of philosophical conceptualisation of a mythical and religious world-view was responsible for the inception of philosophy. Recently Jean-Pierre Vernant (see Vernant 1983) explains the birth of Greek philosophy by an appeal to the development of political and legal practices. Even a very elementary inspection of these theories immediately shows that they are not mutually inconsistent, but rather complementary. Thus, an eclectic position

J. Woleński, *Semantics and Truth*, Logic, Epistemology, and the Unity of Science 45, https://doi.org/10.1007/978-3-030-24536-8_1

seems quite reasonable in looking the phenomenon under discussion. It consists in taking into account of all the mentioned factors and leaving aside the problem as to which of them was the most important or decisive. To some extent, etymological studies on the development of Greek philosophical terminology suggest approaches for general explanations of how Greek philosophical thought developed. Thus, interpretation of concrete linguistic material collected mainly from Homeric poems and other archaic literary sources very essentially depends on whether the given author accepts one of general theories of the development of Greek thought just mentioned or not. Fortunately for persons like myself, there are also numerous studies that focus on more or less hard linguistic data. This circumstance gives an opportunity for some generalizations and comparisons which are, at least partially independent of general account s of the "essence" of Greek civilization.

(**DG2**) I am not an expert in classics. My remarks in this chapter rely heavily on the material which I found in the writings of other people. The most important works I employed (as far as Greek is concerned (references to other languages are provided at the relevant spots) are Levet 1976 and Komorowska 1979. Both books, more linguistic than philosophical, contain a very rich material (I cannot take into account all the linguistic facts registered, including those from the writings by other authors) and provide a very reliable starting point for further philosophical analysis. Of course, my report is also governed by some principles (perhaps even prejudices). I try to find compromises between competing interpretations, using as a guide the rule, already mentioned, that eclecticism is sometimes sound. In general, I rely much more on linguists than philosophers, and I regard etymological studies on philosophical terminology as indispensable for understanding concepts, also philosophical ones. There is a quite extensive (263 p.) book, namely Herbertz 1913, devoted entirely to the concept of truth in ancient Greek philosophy. This book completely neglects etymological and linguistic problems and merely embeds the questions discussed into very general philosophical views (naive realism, idealism, scepticism, rationalism, etc.). I do not deny that the history of philosophical concepts should be closely related to philosophical standpoints, but I do claim that it must be supplemented by a very solid linguistic knowledge. It is particularly important for us to have at our disposal explicit definitions of relevant terms, and their meanings must be extracted from contexts and various usages. The archaic concept of truth is in this very situation. In fact, the first statements that can be regarded as attempts to give definitions of truth appear no earlier than in Plato. My general view is that the philosophical usage of terms (in the present case, the counterparts of 'truth' in archaic and ancient Greek) was related to the archaic one. If we assume, as I do, that philosophy and science is a continuation, at least in some respects, of ordinary life and ordinary ways of thinking, then it is not without interest to investigate which theory of truth accords with a natural development of philosophical terminology, that is, with its transition from the pre-philosophical stage to philosophical standards. I will argue that **STT** employs the core of the concept of truth that is essentially rooted in the archaic Greek language, as well as in other pre-philosophical systems of communication.►

Aletheia is the most important Greek counterpart of our 'truth'; *alethes* (true), *alethos* (truly) and *alethein* (to speak the truth) are related words. However, the Greek truth-family is much more comprehensive and consists of 14 words, among others (adjectives): *atrekes*, *nemertes*, *adolos*, *ortos*, *apseudos*, *etymos* and *etetymos*. It is characteristic that several words, including *aletheia*, belonging to this variety begin with the prefix *a*. The most common reading this lexical phenomenon is to consider this prefix as the sign of a *privativum*—that is, a negative noun or adjective. This understanding of *aletheia* was proposed in antiquity by Sextus Empiricus, Plutarch, Olimpiodoros, and in the so called *Lexicon Gudianum* (see Luther 1935, pp. 12–13, Friendländer 1954, pp. 222, 375). In our times, it was recalled by Leo Myers in his influential *Handbuch der griechischen Etymologie* (1901) and popularized by Rudolf Bultmann (see Bultmann 1933, p. 239): "ἀλήθεια – etymologisch das *Nicht(s)-verheimlichen* – bedeutet". According to this interpretation, we should consider such words as complexes of the following structure: *a-letheia*, *a-trekes*, *a-dolos* or *a-pseudos*; also *nemertes* can be understood in a similar way, because *ne* functions as *a*, namely as the indicator of the privative character, that is, as pointing out a negative form of nouns or adjectives. Logically speaking, privatives have nominal negation (as in 'unbelief'), not always reducible to sentential denial ('it is not the case that'). The etymology of *aletheia* reflects being derived from *a* + *lethe* + suffix. *Aletheia* as a noun occurred in conjunction so-called *verba dicendi*, that is, verbs like the Greek counterparts of 'to tell', 'to say', 'to think' or 'to hear'. So much about matters of lexicology (lexicography) and a very simple grammar. Of course, semantic matters are much more important. Very schematically, the form *V(aletheia)*, where the letter *V* stands for a *verbum dicendi*, represents an *aletheia*-context. An *aletheia* consisted in issuing a concrete sentence about something in the present tense, usually supported by direct experience, particularly seeing (see Boeder 1959, Szaif 1996, pp. 68–71). Then, application of *aletheia*-contexts was extended to past and future events. Finally, *aletheia* became an abstract noun, denoting a property of sentences (judgements, etc.); examples documenting this development will be given in the sequel.

(**DG3**) Many discussions about *aletheia* are strongly influenced by Heidegger's philosophy of truth (see Heidegger 1940). I will not enter into the details of Heidegger's theory of truth, and restrict my remarks to considerations about the archaic and Pre-Socratic meaning of *aletheia*. Heidegger agrees that this word plays the role of a *privativum*. According to him, *a-letheia* principally means *Unveborgenheit*, "disclosure" or "un-concealedness" and stands in opposition to *Verborgenheit*, "closure" or "concealedness"; a more colloquial reading, namely —"which is not hidden" is proposed in Friedländer 1954, p. 221. Heidegger derives his interpretations from the analysis of the famous allegory of the cave in Plato's *Republic*. The fragment 515 in which the word *alethes* (the truth, the true) occurs is crucial. The English translation runs as follows (Plato, *Republic* 515c, tr. by G. A. M. Grube, rev. by C. D. C. Reve, in Plato, *Complete Works*, ed. by J. M. Cooper, and D. S. Hutchison. Indianapolis: Hackett 1997):

> Then the prisoners would in every way believe that the truth is nothing other than the shadows of those artifacts.

Heidegger translates this word as *unverborgene*, that is, disclosed or unconcealed and agrees that *aletheia* is a *privativum*. It is clear that the linguistic material used by Heidegger is extremely poor and is limited to two occurrences of the crucial word; second is in the fragment 515d (my italics):

> […] if we pointed to each of the things passing by, asked him what each of them is, and compelled him to answer, don' t you think he'd be at a loss and he'd believe that the things he saw earlier were *truer* than the one he was now being shown?

Thus, although Heidegger followed the canonical approach to the etymology of *aletheia*, he was guided in his analysis mainly by personal definite intuitions concerning the concept of truth. Heidegger derived very far-reaching consequences from his reading of Plato. He argued that the Pre-Socratic philosophers had only the ontological concept of truth which meant the disclosure of being; according to Heidegger, the further course of philosophy changed this sound understanding and introduced the epistemological concept of truth as something that resides in the mind. The Heideggerian interpretation of *aletheia* was strongly criticized by Paul Friedländer (see Friedländer 1954, pp. 221–229). He pointed out that Heidegger overlooked the ambiguity of *aletheia* which in Plato meant reality of being or correctness of apprehension and assertion (p. 227). In order to demonstrate Heidegger's errors, Friedländer suggests that the interpretation of *aletheia* as a *privativum* is perhaps not correct—but he does not recommend, however, without outlining a firm alternative. Friedländer was criticized by Wilhelm Luther (see Luther 1966, p. 34f.) who insisted (Luther 1935, pp. 11–12) that the canonical interpretation of *aletheia* was right. He also agrees that the archaic usage of this word was entirely ontological. It seems that Friedländer looked for an alternative reading interpretation of *aletheia* just in order to criticize Heidegger. It is a surprising strategy, particularly if no alternative is even sketched. A more serious argument by Friedländer's is that *aletheia* as a *privativum* occurs only once in the archaic texts (namely in Hesiod's *Theogony*, where the sea god Nereus is described as *a-pseudos* and *alethes*). Luther replies (p. 34) that one can find further examples in Sophocles and Euripides. Friedländer's strongest argument against Heidegger's interpretation of *aletheia* points out that it is semantically inadmissible (p. 223):

> Thus, in the one case where in early times *ἀληθής* is understood as *ἀ-ληθής*, it has nothing to do with the hiddenness of being, but designates a person who does not forget or neglect, who does not lose something out of sight or mind; in short, it means exactly the "correctness of perception" which Heidegger in his sketch of the words *ἀληθής* and *ἀλήθεια* attributes to a period of the decline of Greek thought.

It is clear that this argument holds even if *aletheia* occurred very frequently as a *privativum* in the archaic Greek. Now, it is beyond any doubt (see below) that *aletheia* and cognate words are ambiguous. Friedländer can still be right even if there are no reasons to claim that *alethes* in archaic Greek referred only and exactly to the correctness of perception. In order to reject Heidegger's interpretation it is quite sufficient to demonstrate—what is rather a simple task—that *aletheia* expressed various contents. Heidegger is certainly right that Plato's words translated as 'truth' and 'truer' did not express epistemological relations but properties of being; clearly, Plato had in mind forms as true or truer being. On the other hand, the meaning attributed by Heidegger to the word *aletheia* was rather rare or secondary (see Szaif 1996, pp. 145–146, note 92 for a brief summary of objections against Heidegger). The quoted translation of *Republic* 515c probably renders Plato's intuitions adequately. However, the discussion is difficult, because we encounter in particular interpretations an explicit or implicit intervention of philosophical insights (perhaps even a priori) into etymological analysis. Anyway, it seems that neither Heidegger nor Luther proved that the basic usage of *aletheia* in archaic Greek was purely or even mostly ontological. And this is a crucial point for philosophical issues concerning *aletheis*. In order to complete this digression, I would like to note that the relation of Luther to Heidegger is quite complex. In Luther 1935, p. 12 we read:

> Die Grundbedeutung von *λήθω* [*lethe* – J. W.] liegt in Richtung der deutschen Konzeptionen "ich bin verborgen, verdeckt, verhüllt". […]. Von der ermittelten Grundkonzeption aus erklärt sich ferner die Privatbildung *ἀλήθεια* […]. Ihre etymologische Bedeutung is "*Unverborgenheit*".

Speaking about German conceptions (*deutsche Konzeptionen*), Luther alludes to Heidegger who is mentioned, with a very great respect, few pages earlier. On the other hand, Luther in his later study (see Luther 1966, pp. 172–173) considered Heidegger's interpretation of Plato on truth as completely mistaken. To sum up: for Luther, (a) *aletheia* is a *privativum*; (b) *aletheia* means disclosure, but only in early Greek philosophy (see also the next section); (c) the concept of truth in early Greek philosophy was ontological; (d) Plato was a predecessor of the correspondence theory of truth. Thus, the main of disagreement between Luther and Heidegger was how to interpret Plato, although it is still not quite clear whether Luther's understanding the former understood (particularly, in Luther 1966; also see Luther 1965) of *Unverborgenheit* is the same as Heidegger's did.►

What about the meanings of others words belonging to the truth-family? It is convenient to consider some adjective-forms. *Alethes* can be translated (aside of 'true') by 'unhidden' (also contrasted with 'silent' or 'forgotten'), *atrekes* by 'not deformed', *nemertes* by 'faultless', *adolos* by 'not deceitful', *ortos* by 'simple-minded' (or 'not double-faced'), *apseudos* by 'truthful', *etymos* and *etetymos* by 'real', 'actual' or 'authentic'. Inspection of these meanings (or ambiguities, if you like) immediately shows that words from the truth-family express ontological, epistemological and moral contents, integrated into utterances used in ordinary situations. Particular items from the truth-family could be used, and were used, interchangeably. It is certain that dialogues in Homer's poems were modelled on ordinary dialogical situations in which to say 'it is true' and 'it is actual' means the same; the same situation occurs in our contemporary life. Concrete dialogical exchanges are very sensitive to lying, deceiving, etc. and, thus, truth was contrasted rather with lying than falsehood; it seems that in archaic Greek there was no sharp difference between lying and uttering false sentences. Although the word *pseudos* became the most popular from the Greek falsehood-family, this variety counts 67 items, much more than the truth-family; it is indirect evidence that lying and deceiving impressed the Greeks more than telling the truth. Perhaps we can consider this circumstance as an expression of a natural tendency, also present in the contemporary use of language that regarding a statement as true is a primary attitude, possibly to be corrected by further data.

It is possible that the moral and ontological dimensions of *aletheia* and related words occurred earlier, but their constructions with *verba dicendi* indicated that the epistemological usage was also not absent in the archaic employment of language. Krischer 1965, Snell 1975 and Komornicka 1979 point out that *aletheia* meant a state of affair being real or actual in relation to the epistemic state of the teller or knower. On the other hand, according to these authors, *etymos* is more (or even purely) ontological and refers to fully actual states of affairs, independently of any involvement of epistemic attitudes. Robert Bultmann (Bultmann 1933, p. 239) thinks that *aletheia* connoted that both aspects, but that the ontological functioned as more fundamental. It is not for me to discuss which view is correct. What I can derive from this account is that the epistemic factor early the supplemented ontological one (assuming that the latter was really prior). Thus, we have no reason to deny that 'to tell truth' in archaic Greek also expressed 'to tell as things are'. On the other hand, since there was no advanced philosophy surrounding this language, it is difficult to look for explicit philosophical contents associated with truth-talk in which the archaic *aletheia* was involved—for example, to consider the problem of truth-bearers (although it is possible to maintain that 'to say *aletheia*' was understood as 'to issue a true utterance') or relations of truth to logic (there was no theoretical logic at that time; the adjective 'theoretical' indicates that the question of logic as a natural human capability is not taken into account in my comments).

(**DG4**) We do not know who originally formulated the Liar Paradox and in which wording. It is usually reproduced by the following puzzle. Epimenides, the Cretean, says 'I am just lying'. Is his statement true? If it is true, it is false, because he is not telling the truth, but if it is false, he is telling the truth. Diogenes Laertius (*Lives of Emiment Philosophers*, tr. by R. D. Hicks, Harvard University Press, Cambridge, Mass. 1925, vol. 2, p. 108) attributes it to Eubulides of Megara (see also Rüstow 1910, Cavini 1993). Diogenes Laertius seems to be the first who used the label 'the Liar' for this paradox, which is semantically, the antinomy arising from self-referential use of the predicate 'is false'. Cavini conjectures (p. 99) that "the Greek Liar […] just says: '*ἐγώ ψεύδομαι*' [it is false – J. W.]." That the paradox in question was baptized 'the Liar Paradox' seems to confirm the view that there was a lack of sharp distinction between lying and telling falsehoods. See also the next digression and Epidemides' fragment quoted there.►

(**DG5**) Three texts from the transition of the archaic to the philosophical period are fairly interesting: (A) Epimenides, the Cretean (Diels I, B1 3): “Creteans always lie” (*profasis*, *pseudai*); (B) Solon (Diels I, 10β 6: “Don’t lie, tell truths” (*ne pseudos*, *all’ aletheie*); (C) Pittakos (Diels I, 10ε 12): “Take care about piety, […], common sense, truth (*aletheias*), […], friendship.” These are very instructive fragments. Epimenides accuses Creteans of always deceiving; however, (B) is not equivalent to the Liar Paradox. Pittakos locates truth among other personal virtues. Both fragments clearly point out the moral dimension of telling truth and lying. One can conclude that since Solon was the legendary legal reformer, his words apply to legal matters (perhaps “Don’t lie before the court”). On the other hand, several of Solon’s preserved maxims have nothing to do with law. Since all can be interpreted as principles of practical wisdom, one can say that also (C) offers a moral prescription. Moreover, the second part of (C) possesses the crystal grammatical structure consisting of *verbum dicendi* (*all*) and *aletheia*. It is interesting that the ontological interpretation of (C) seems to be quite artificial, while the epistemological (tell about things as they are) looks fairly natural.►

(**DG6**) In order to complete somehow the above remarks, I finally sketch an interpretation of the archaic *aletheia* developed by Marcel Detienne (see Detienne 1967). He, following Vernant, looks for a general social context in which truth-talk arose. In particular, Detienne points out that habits concerning *aletheia* were formed by religion (in fact, *Aletheia* had a divine personification as a daughter of Zeus), myths, poetry, and legal (truth as justice) and political practices. Detienne (p. 49) constructs two sequences. One (positive) consists of Praise, Speech, Light, Memory, and *Aletheia*; second (negative) includes: Blame, Silence, Darkness, Oblivion, and *Lethe*. Detienne argue that there was no sharp opposition (contradiction) between *positiva* and *negativa* (the same was noted by Heitsch 1962, p. 31), but elements of the particular parts were partly contrastive and partly complementary; they form, according to Detienne, a semantic field (a category of structural linguistics). I will not discuss this topic, because I do not see any special benefit from applying a very controversial linguistic theory to our problem; similarly, the hypothesis about pre-logical thinking explains very little, if anything at all. It is sufficient to note that *aletheia* and *pseudos* were regarded as different, at least in the time of the Seven Sages (see (C) in **DG5** above). As far as other points are concerned, note that I choose only one fragment of Detienne’s book (he presents several other tables with two opposite-complementary sequences). Of course, Detienne collects well known facts, for example, the connection of light and truth. However, I think that Detienne exaggerates in his insistence about the connection between truth, memory and tradition. Certainly, there are reasons for understanding ‘true’ as ‘present in memory or tradition’ or ‘false’ as ‘forgotten’, but it is worth noting that, due to the lack of other records, memory and tradition functioned as sources of knowledge or conditions of the possibility of telling the truth—particularly about the past. Similarly, the prophetic power of *aletheia*, also associated by Detienne with its archaic uses, had an obvious link with the already mentioned extension of applying truth-talk to the future. Thus, there is nothing in the facts pointed out by Detienne that would compel us to abandon the interpretation of the archaic *aletheia* as telling as things are.►

(**DG7**) Although Greek is crucial for studies of the development of philosophical terminology, it is interesting to consult how truth-talk functioned in other languages. Hjalmar Frisk (see Frisk 1936) collected various facts concerning Indo-Germanic languages. Although his study focuses on syntactic and morphological aspects, it also provides some semantic information. It confirms, for instance, that several words translatable as 'truth' or 'true' were opposite to expressions associated with lying, and that truth-words were ambiguous in a similar way as in Greek. In particular, truth-words refer to reality, truthfulness, justice, law (Polish 'prawda'), certainty, etc. The genesis of the word *Wahrheit* (truth) in old German also indicates (personal communication of Aleksander Szulc) various aspects and ambiguities. The word *wēr* (to defend, to guard) is considered as the core of the later Germanic counterparts, like *wēra*, *wār* and, finally, *wahr*. Other cognate words *wērō* and *wārō* refer to grace, obligation, contract, consideration or truth in its various applications. Further, *wār-haft* meant 'what is reliable'. These remark show how strong were moral and legal factors of truth-talk were in old German.►

(**DG8**) The concept of truth was studied by many Chinese philosophers (see McLeod 2016). I will mention two cases relevant to my main historical issue. *Lunyu* (Confucian Dialogues) two words *ran* and *shi* are related to naming and truth (Chinese philosophers usually associated assigning a proper name with telling truth. When Confucius asks one of his students (McLeod 2016, p. 46) "Given this is so, is *Shi* superior?" Alexus McLeod argues (p. 52) that various passages in *Lunyu* seem

> to recognize at least something like a correspondence intuition […], a notion that proper (assertoric) speech is that which is grounded in reality or fits "the way things are".

Xu Gan, a 2nd century C.E. philosopher, affords another example of correspondence thinking (McLeod 2016, p. 164):

> Xu uses *ran* […], *shi* […], and related terms in ways that clearly suggest truth in any recognizable form by us or anyone else. […]. Just as in the *Zhuangzi* and *Huainanzi* [both are compilations of philosophical texts – J.W), proper statements are those that correspond somehow to *dao* [way, reality] […] properly used names are those corresponding with actuality.

(**DG9**) In Hindu and Buddhist traditions, there is very close connection (see Vroom 1989, Chaps. III–IV) between truth (*satya*, *bhutam*, *Dharma*) and religion. However, we can discover explicit or at least implicit epistemological points in related doctrines. In general, both traditions distinguish truth on the world (things) and its knowledge. Although Buddhism is closer to scepticism than Hinduism, both systems insist that knowledge is valuable and has to do with the truth of things conceived (Vroom pp. 131, 171) "as things are" (ontological truth). Some authors (see Kumari 1987) claim that Buddhist aletheiology can be considered as a version of pragmatism.►

(**DG10**) Ancient Egyptian thought connected truth with politics and religion. The word *maat* had many meanings and it also referred to truth as the religion of kings and was opposed to lying as characteristic beliefs of ordinary people (see Assmann 1990, p. 232). However, an epistemic importance of truth can be easily extracted from this contrast. See also Van De Mieroop 2015 for the analysis of similar fact about the ancient Babylonia.►

(**DG11**) We also have studies (Kittel 1933, pp. 233–238, Michel 1968, Vroom 1989, Chap. V, Wójcik 2010) about *emet*, a Hebrew word as used in the Old Testament and Talmud. Its opposite is *sheker*, used for lying as well as for falsehood. *Emet* (Quell 1933, p. 233)

> ist verwendet als Bezeichnung einer Wirklichkeit, die als [..] *fest*, daher *tragfähig*, *gültig*, *verbindlich* anzusehen ist und somit *Wahrheit* bedeutet. (is used to designate reality, which is to be considered as established, ready, valid, binding and therewith means truth).

Quell's account is parallel to that of Bultmann concerning many uses of *aletheia* and is confirmed by explanations of Saadia Gaon (Saadia ben Josef; 882 or 892–942) in his book *The Book of Beliefs and Opinions*, New Haven: Yale University Press 1948. He attributes to *emet* (p. 471) the following meanings (I omit some): (a) attained by study; (b) sometimes not recognized; (c) can be mistaken for (falsehood); (d) duty to direct man to; (e) of things perceived without illusions; (f) ways to refuting scepticism; (g) science, knowledge and means to achieve them; (h) assertion about a thing what it is. Similar points are also noted by Hendrik Vroom (p. 197). Diethelm Michel (pp. 38–40) particularly focuses on two applications of *emet*: (a) as correctness of sentences, and (b) as moral or legal rightness. As the matter concerns (a), Michel concludes (p. 40) that

> ämät bezeichnet in den am häufigsten vorkommenden Wortbedingungen, die deshalb als typisch anzusehen sind, das Übereinstimmung einer Aussage mit einer Tatbestand.►

(**DG12**) The recent and fast growing interest in African philosophy also resulted in some observations about the concept of truth as used in this cultural environment. Here is an overall description (Bello 2004, p. 271):

> […] it has been shown that the word *iro* in Yoruba, like the word *nkontompo* in Akan, has primarily a moral meaning, since it means "lie." *Iro* (*nkontompo*) is the opposite of *notito* (*nokware*), which means "truth". But the opposite of "truth" in English is "falsehood," not "lie". This means that both the Yoruba and Akan languages have affinities in the matter of the concept of truth and its cognates.

However (see above about archaic Greek, old German (see **DG7**), and on *emet* and *sheker* (see **DG10**) the mentioned affinities were more extensive. A more epistemological meaning of truth is recorded by Kwasi Wiredu, a contemporary African philosopher from Ghana (quoted after Hallen 2004, p. 107):

> To say that something is true, the Acan simply says that it is so and so, and truth is rendered as what is so.

Even if Wiredu uses the current philosophical intuitions, he certainly reproduces something deeply rooted traditional African culture.►

(DG13) Although various languages functioned in different contexts and expressed very different meanings, frequently very different from those associated with philosophical issues, there is definitely something that they share with respect to the uses of counterparts of 'truth' and 'true'. Thus, we can assume that when philosophers became to be interested in the nature of truth as (speaking in our contemporary jargon) correspondence with facts, they followed a definite linguistic path present in many (perhaps even in all) languages. Yet there is the significant question as to, whether the archaic discourse about truth should be qualified as ontological or epistemological (I anticipate a controversy, which will be dealt with in Chap. 4). Although reference made to "how things are", looks ontological, the entire context alluding to knowledge and its properties, also exhibits epistemological features. Even if we say that both aspects, ontological and epistemological, are conflated, the latter is ruled out.►

(DG14) In **DG2** I critically mentioned Herbertz 1913. This monograph tries to outline truth-theories of several "silent" (that is, not speaking about truth) philosophers. For example, Herbertz derives the theory of truth of the Miletians from their naive epistemological realism. We encounter more or less similar attempts in interpretations reported in some previous digressions. Take, for instance, Michel's remarks on *emet* reported in **DG11**. Michel's conclusion could very please philosophers who, like me, are looking for the historical roots of the idea of correspondence (agreement) with facts (according to the formula expressed by the phrase *das Übereinstimmung einer Aussage mit einer Tatbestand*) as the base of truth-talk. However, I think that Michel simply applied a popular contemporary philosophical label in order to explain linguistic facts from a more or less remote past. Examples taken by Michel from the Old Testament (for instance, the requirement that truth should be said) actually allow us to say that *emet* was regarded as a property of sentences or words, but we have no evidence for use of the word 'correspondence' between sentences and facts. Similarly, McLeod's reference to the concept of correspondence is perhaps too burdened by his access to present philosophical jargon. All that can be derived from the examples and their description is that *shi*, *ran*, *emet*, *notito* or *nokware* were employed used for telling how things are.►

1.2 Pre-Socratics

No preserved fragment of the philosophers from Miletus speaks about truth by using *aletheia* or any other word from the Greek truth-family. I do not know how to explain this fact. The Miletians' interest in the philosophy of nature provides no hint, because this style of doing philosophy does not preclude thinking about truth. Likewise, nothing is achieved by pointing out that the relevant writings may have disappeared because if had the Miletians spoken about truth, it would be recorded by secondary sources. For instance, Sextus Empiricus (*Against the Logicians*, Book I, tr. by R. G. Bury, in *Against the Professors*, vol. 1, Cambridge, Mass: Harvard University Press 1925) reports many views about truth but does not mention the Miletians in this context. Incidentally, the silence of the first philosophers of nature about truth should be particularly astonishing for the adherents of the ontological understanding of *aletheia*, because if this word primarily refers to the world, it should have been used by the Miletians. Still more surprising is that *aletheia* does not occur in the Pythagoreans despite of their interest in mathematics, ethics, and, in particular, the practice of doing very advanced deductive inferences.

Historically, Xenophanes of Colophon and Heraclitus seem to be first philosophers to speak about truth, at least according to existing sources. The former did it that in two preserved fragments (Kirk, Raven, Schofield, p. 179):

(1)

No man knows, or ever will know, the truth about the gods and about everything I speak of; for even if one chanced to say the complete truth, yet oneself knows it no; but seeming is wrought over all things;

(2)

Let these things be opined as resembling the truth.

The first fragment contains a word *tetelesmenon*, also, although seldom, used in archaic Greek for truth, but the second employs *etymos*. The word *aletheia* occurs once in Heraclitus (Diels I 22B 112; Eng. tr. in D. W. Graham, *The Texts of Early Greek Philosophy. The Complete Fragments and Selected Testimonies of the Major Presocratics.* Part I, Cambridge: Cambridge University Press 2010, p. 171:

Sound thinking is the greatest virtue and wisdom: to speak the truth and to act on the basis of an understanding of the nature of things.

And this is all that can be found in writings of the older philosophers of nature (including Heraclitus), the Pythagoreans, and the first of the Eleatians.

(**DG15**) The lack of the words belonging to Greek truth-family in the Miletians, the Pythagoreans and also so called younger philosophers of nature (Empedocles, Anaxagoras) did not use the words belonging to Greek truth-family, does not preclude attempts to reconstruct concepts or even theories of truth presumably maintained by these philosophers, because the lack of definite words does not mean the lack of related concepts. Such reconstructions are based on general views held by particulars thinkers and their remarks about related questions, for instance, ways of cognition and began to be practised even in ancient times. For example, Sextus Empiricus (*Against the Logicians* I, 115) informs that Empedocles gave six truth-criteria; it is considerations concerning methods of knowledge-acquisition that gave rise to various later theories of truth-criteria. The most advanced historical analysis of this kind is to be found in Luther 1966. He uses not only the general views of particular philosophers, but also their concrete statements. For instance, Luther observes that something follows for the problems of truth from remarks of Alcmaeon of Croton about clarity and exactness of cognition; from Heraclitus' ideas of *logos* and reason as well as from his explicit remarks about the fallibility of senses, from Empedocles theory of experience; from Anaxagoras' account of *nous* or from Democritus' view that there are two different modes of cognition: obscure (via senses) and true (via intellect) (see Diels II 68B 9–11; Kirk, Raven and Schofield, pp. 409–410) (it is interesting to note that Democritus used the truth-talk in a more abstract way than it was done earlier). We encounter here a very important feature of passing from the stage of pre-philosophy to the stage of genuine philosophy. The latter is associated with the existence of a web of concepts. It enables historians to fill gaps in a given web, for example, *logos*, *nous* and properties of sense-experience. Although I appreciate these attempts, I am mainly interested in explicit statements about truth, but these are possible only if concrete words are used. This orientation dictated that I will not enter, except the next digression, into the fairly impressive reconstructions by Luther. On the other hand, to assume that *aletheia* (or its equivalents) began to function in the abstract conceptual apparatus of Pre-Socratic philosophy seems to be an admissible hypothesis.►

(**DG16**) Luther reconstructs the truth-theories of the Pre-Socratic philosophers as ontological. It is even surprising in some cases, because he often derives his conclusions from typical epistemological categories, like clarity, exactness, logos, reason or sense-experience and its various properties. Thus, Luther extends his account of *aletheia* to views expressed in a new vocabulary. I believe that one point requires some critical comments. Luther argues that early Greek historians, notably Herodotus, understood truth ontologically. He reports (p. 88) that Hekataios of Miletus, a historian (one generation before Herodotus, that is, about 500 B.C.), used the word *aletheia* in the description of his way of writing history. He says that he will write what the truth (*aletheia*) is, according to his experience; thus, his use was more abstract than the ordinary one, but still coloured by personal experience. Herodotus, who is regarded as the first methodologist of history, also speaks of true (*alethes*) reports about the past. Charles Kahn (see Kahn 1973, *passim*) quotes several veridical uses of *to eon* (to be) in Herodot, for instance (pp. 352–353):

> He should not have told the truth [*legein to eon*], if he wanted to lead the Spartans on an expedition into Asia.
>
> Cleomanes asked Crius his name, and the latter told him what is was [*to eon*] (told him truth).

According to Luther, Hekataios and Herodotus (by the way, Herodotus' examples show how *aletheia*-contexts were employed in concrete cases) claimed that historical truth was the task of writing history based on the best of experience. Yet he (Luther) insists that both Greek historians understood *aletheia* (and its substitutes) as an ontological category. Is it not simpler to say that methodologically self-conscious historians were interested in telling how things were? We have also a very interesting fragment in Thucydides (*History of the Peloponnesian War* I, 2, tr. by C. Foster Smith, Heinemann, London 1919):

> So averse to taking pains are most men in the search for the truth [*aletheias*], and so prone are they to turn to what lies ready at the hand.

I do not know whether Luther would be ready to include Thucydides (he was a contemporary of Socrates) among the ontologists in truth-theory. It is clear that the replacement of 'truth' by 'reality' transforms the quoted fragment into a text that is not quite coherent internally. This proves that Thucydides understood *aletheia* in an epistemological manner; moreover his understanding is abstract and intersubjective. In fact, he commented on several obstacles to apprehending historical truth, that is, to a correct report on how things were. Moreover, we should rather look at the link from Hekatoias and Herodotus, and Thucydides, which is much more comprehensible, than at the engagement of former historians with the concept of disclosure as the meaning of *aletheia*. I do not insist that locutions involving *aletheia* and *alethes* in Hekatoios and Herodotus are completely devoid of ontological aspects, but I think that Luther's interpretations are too one-sided as a result of his philosophical biases. If we inspect Kahn's translations (see above), he assumes that telling the truth consists in telling how things are (were, will be). It supports the view that truth-talk in archaic Greek was not limited to the ontological dimension of *aletheia* and its equivalents (see remarks at the end of the previous section).►

Since I consider the problem broached in the last digression as rather crucial, the following comments are in order. Even if we say that *aletheia* concerns things or being, or is about things or being, that is not sufficient for the ontological conception of truth. It requires considering truth as a feature (I do not enter here into the problem of the meaning of 'feature' in this context) of things or being. Plato did that in the *Republic* (see **DG3**), but there is no convincing evidence that this view on truth was prevailing. Of course, I do not deny that some uses of *aletheia* in the Pre-Socratics were ontological. Consider for example, the following fragment of Democritus (more precisely, it is a report by Sextus Empiricus about Democritus' views; Kirk, Raven and Schofield, p. 410):

> Democritus sometimes does away with what appears to the senses, and says that that none of these appears according to truth [*aletheias*] but only according to opinion: the truth [*alethes*] in real things is that there are atoms and void.

Clearly, *alethes* refers to real things and *aletheias* can be replaced by 'reality' or 'actuality'. On the other hand, truth is here contrasted with opinion (*doxas*)—which also invites an epistemological reading.

Heraclitus' mentioned passage about truth does not bring anything particularly new; that wisdom consists in saying the truth can be implicitly found in Solon or Pittakos (see **DG5**). But Xenophanes' ideas actually began a new chapter in the history of epistemology. In fact, the fragment "Let these things be opined as resembling truth" introduces a distinction of the utmost importance, namely, that between opinion and truth, although in Xenophanes' case, it is exclusively restricted to theological matters. This distinction became one of the main themes of Parmenides' thought (and the majority of philosophers until now). There are three important fragments about truth in Parmenides (Kirk, Raven and Schofield, p. 245, pp. 254–255):

(a)

Come now and, and I will tell you (and you must carry my account away with you when you have heard it) the only ways of enquiry that are to be thought of. The one, that [it] is and that it is impossible for [it] not to be, it the path of Persuasion (for she attends upon Truth [*Aletein*]); the other, that [it] is not and that it is needful that [it] not be, that I declare to you is an altogether indiscernible track; for you could not know what is not – that cannot be done – nor indicate it.

(b)

Here I end my trustworthy discourse and thought concerning truth [*aletheies*]; henceforth learn the beliefs of mortal men, listening to the deceitful ordering of my words.

(c)

It is proper that you should learn all things, both the unshaken heart of well-rounded truth [*Aletheies*], and the opinions of mortals, in which there is no true [*alethes*] reliance. But nonetheless you shall learn these things too, how what is believed would have to be assuredly, pervading all things throughout.

The fragments (b) and (c) introduce the distinction between truth and opinion in full generality—not only restricted to a particular field, as in Xenophanes. It is customary to render this distinction as the contrast between *episteme* and *doxa*. Thus, truth is a property of *episteme*, or we can even say that the domain of truth and the domain of *episteme* are identical.

Let me simply register this function of truth without further comments, although it merits them—but perhaps not in the context of the concept of truth; the history of epistemology from Plato to Gettier's counterexamples documents that the trouble is with defining *episteme* (knowledge) rather than truth (however, see Chap. 9, Sect. 9.10). Thus, there remains a host of problems connected with (a)—for example, the question how truth is related to being and existence. It is well-known that there are many difficult issues here enhanced, additionally strengthened by the cryptic Parmenides thesis that thought and being are the same—and many proposals for to solve the difficulties in interpreting Parmenides (see Wiesner 1996 for an extensive discussion). I cannot enter into this interpretative jungle. Fortunately, my task justifies restriction to a few rather simple remarks. According to Parmenides, truth is extensionally equivalent to being. It justifies the veridical use of '*to on*" (to be) (see Kahn 1973, Chap. VII for further discussion). Thus, to say '… is' and '… is true' appear sometimes (of course, not always) as equivalent. *Aletheia* in this sense is really an ontological category, *pace* Luther, but it is clear that this understanding is only one of those possible. Even (a) can be interpreted more epistemologically if one will focus more on "enquiry" than on "that [it] is and that it is impossible for [it] not to be", but the most plausible interpretation of Parmenides is that he focused on being. Incidentally, the equivalence of truth and being plagues (I do not exaggerate) philosophy from the Eleatians to our days. The problem concerns the truth of negative existentials: if truth and being are equivalent, it is mysterious how it is possible to establish truthful sentences about objects that do not exist, because non-existing objects, due to the assumed equivalence, seem to lead to an obvious inconsistency (see Pelletier 1990, Perzanowski 1996 for a discussion of this problem). Plato was the first who tried to resolve this difficulty (see below).

For Parmenides, the contrast between being and non-being is sharp and exhaustive: the product of being and non-being is empty, and there is nothing in between being and non-being. If we map these principles pertaining to truth onto logic, we receive the laws of non-contradiction and the excluded middle—regardless whether they are understood ontologically or logically. In any case, if it were true, as Heitsch and Detienne maintain (see **DG6**), that *aletheia* and *pseudos* were at least partly complementary in archaic Greek that passed away with Parmenides. This is not surprising if we remember that the Eleatians contributed essentially to the rise of logic, perhaps not logical theory (this had to wait for Aristotle), but to logical deductive practice. The last question is whether the idea of correspondence between thought and reality was involved in Parmenides' thinking about truth. The answer should be "yes" but with various qualifications. Certainly, not in the sense usually attributed to Aristotle or Thomas Aquinas. On the other hand, the extensional equivalence of truth and being (as well as the identity of being and thought, whatever that might mean) strongly suggested that truth is a relational concept—because truth-talk is related to being. Thus, the novelties and problems introduced by Parmenides to the philosophy of truth were numerous and fairly fundamental, including his popularization of the word *aletheia* as the standard linguistic label in referring to truth. The role of Parmenides is well summarized in the following quotation (Peters 1967, pp. 16–17):

> The presence and even the possibility of truth are closely related to the Greek distinction between *doxa* and *episteme* […] and their proper objects. Thus there is really no critical problem until Parmenides distinguishes being from nonbeing, associates the later with sense perception, asserts that there is no truth in the phenomenal world of *doxa* […], and contrasts the latter with the "way of Truth" […]. As a corollary of this and of the realization of the arbitrary nature of laws and customs […] Protagoras propounded his theory of the relativity of truth […].

Protagoras was presumably the author of the first work whose title contains the word *aletheia*; according to the tradition, he wrote a treatise *Aletheia e kataballountes* (Truth or Refuting Speeches). The only fragment preserved. It is the famous passage in which the doctrine of man-the-mensura is expressed (quotation after Guthtrie 1969, p. 183):

> Man is the measure of all things that are that they are, and of things that are not that they are not.

Typically, the first part of this statement is quoted and taken as the formulation of Protagoras' relativism and subjectivism: if man is the measure of all things, then everything is relative or subjective. However, Protagoras, in order to develop his anthropological views, did not need to add the second part of his maxim; he could rely on practical and rhetorical premises.

The title of Protagoras' work suggests that he wanted to refute some views about truth. It is clear that Parmenides' absolutism and rationalism was Protagoras' target. The sequel to 'Man is the measure of all things', that is, the phrase '[of things] that are that they are, and of things that are not that they are not' is formulated in the Parmenidean language. The ending of the phrase—'things that are not that they are not'—clearly contradicts Parmenides, who rejected the idea that there are things that are not. The first part, namely '[of things] that are that they are', very clearly expresses Parmenides' approach to truth. Protagoras rejected this understanding of truth by admitting that human opinions decide of things that are that they are and of things that are not that they are not, and thereby dictate truth. This view devastated logic, which, as I pointed out above, came into its existence with the Eleatians. In fact, Protagoras maintained that everything is true, Gorgias that everything is false, and both views destroyed the principle of contradiction as it happens in the case of every one-valued logic (this statement uses contemporary metalogic). I do not deny that the discussions of Socrates, Plato and Aristotle with Protagoras and other Sophists had enormous historical significance. However, these debates focused (and still do; see Classen 1989) mainly on the first part of his statement, and thereby concealed an important matter which seems relevant for the history of the concept of truth. Protagoras in his man-the-mensura maxim introduced a canonical language for speaking about truth, although it is not clear whether sentences or sensations were for him the primary bearers of truth (we have the same ambiguity is in Democritus). If we add 'to tell truth is to tell' to 'of all things that are that they are and of things that are not that they are not', we obtain a prototype of the famous formulas of Plato and Aristotle (see Kahn 1973, p. 367) that certainly do apply to sentences. Thus, Protagoras is "a corollary" to Parmenides not only for replacing absolutism by relativism, but also by developing a suitable language in which "The Way of Truth" could be expressed without poetic allegories. Another important point is that Protagoras who was a typical epistemologist, probably contributed to the rise of the view that truth is in intellect, and not in things. If I am right, then, surprisingly, Protagoras—who refuted Parmenides and was in turn refuted by Socrates, Plato and Aristotle—became a link between the Eleatians and the Big Three philosophers from Athens.

1.3 Plato

Plato developed the idea of *episteme* and made it one of the most crucial points of his philosophy. According to Plato, *episteme* has its own special subject matter: the world of forms. Thus, *episteme* is the true knowledge about the true being. It is the ontological doctrine of truth. Some authors (see Szaif 1996, p. 15) speak about ontological-epistemological theory of truth in Plato, but I avoid this qualification because I reserve the label 'epistemological theory of truth' for views that consider truth as a property of products of cognitive acts, although I do not deny that Plato's ontological theory of truth was not restricted to ontology and had an explicit epistemological dimension. This section is exclusively devoted to Plato's definition of true and false sentences (propositions); I leave apart Plato's account in *Cratylus* where he speaks about truth of names.

Plato's definition of propositional truth is given in the *Sophist.* There are relevant fragments in two translations (I deliberately quote these rather long fragments it order to show some details of Plato's thinking on truth):

I. *Sophist*, tr. by N. P. White, in Plato, *Complete Works* (see **DG3**):
(261e–263b):

> Visitor [of Elea]: […] there are two ways of use your voice of indicating something about being.
>
> Theaetetus: What they are?
>
> Visitor: One kind is called names, and the other is called verbs.
>
> Theaetetus: Tell me what each of them is.
>
> Visitor: A verb is the sort of indication that's applied to an action.
>
> Theaetetus: Yes.
>
> Visitor: And a name is the kind of spoken sign that's applied to things that perform the actions.
>
> Theaetetus: Definitely.
>
> Visitor: So no speech is formed just from names spoken in a row, and also not from verbs that are spoken without names.
>
> Theaeteus: I didn't understand that.
>
> Visitor: Clearly you were focusing on something else when you agreed with me just now. What I meant was simply this: things don't form speech if they're said in a row like this.
>
> Theaeteus: Like what?
>
> Visitor: For example "walks runs sleep", and other verbs that signify actions. Even if somebody said all of them one after another that wouldn't be speech.
>
> Theaeteus: Of course not.
>
> Visitor: Again, of somebody said "lion stag horse" and whatever names there are of things that performs actions, the series wouldn't make up speech. The sounds he uttered in the first and second way wouldn't indicate either action or inaction or the being of something that is not something that is or of something that is not – not until he mixed verbs with nouns. But when he did that, they'd fit together and speech – the simplest and smallest kind of speech, I suppose – would arise from that first weaving of name and verb together.
>
> Theaeteus: What do you mean?
>
> Visitor: When someone says "man learns", would you say that's the shortest and simplest kind of speech?
>
> Theaeteus: Yes.
>
> Visitor: Since gives an indication about what is, or comes to be, or has come to be, or is going to be. And he doesn't just name, but *accomplishes* something, by weaving verbs with names. That's way we said that he speaks and doesn't just name. In fact this weaving is what we use the word "speech" for.

Theaeteus: Right.

Visitor: So some things fit together and some don't. Likewise some vocal signs don't fit together, but the ones produce speech.

Theaeteus: Absolutely.

Visitor: But there's still this one point.

Theaeteus: What?

*Visitor: Whenever there's speech it has to be about something. It's impossible for it not to be about something.

*Theaeteus: Yes.

Visitor: And speech also has to have some particular quality.

Theaeteus: Of course.

Visitor: Now let's turn our attention to ourselves.

Theaeteus: All right.

Visitor: I'll produce some speech by putting a thing together with an action by means of a name and a verb. You have to tell me what it's about.

Theaeteus: I'll do it as well as I can.

*Visitor: "Theaeteus sits" That's not a long piece of speech, is it?

*Theaeteus: No, not too long.

*Visitor: Your job is to tell me what it's about, what it's of.

*Theaeteus: Clearly it's about me, of me.

*Visitor: Then what about this one?

*Theaeteus: What one?

*Visitor: "Theaeteus (to whom I'm now talking) flies."

*Theaeteus: No one could ever deny that it's of me and about me.

*Visitor: We also say that each piece of speech has to have the same particular quality.

*Theaeteus: Yes.

*Visitor: What quality should we say each one of these has?

*Theaeteus: The second one is false, I suppose, and the other one is true.

*Visitor: And the true one says *those that are*, as they are, about you.

*Theaeteus: Of course.

*Visitor: And the false one says things different from *those they are*.

*Theaeteus: Yes.

*Visitor: So it says *those that are not*, but that they are.

*Theaeteus: I suppose so.

(II) *Sophist*, tr. by F. M. Conford, in Conford 1935, pp. 308–312; I give only those counterparts of above the fragments indicated by (*):

Str[anger]: Whenever there is statement, it must be about something; it cannot be about nothing.

Theaet[eus]: That is so.

Str. "Theaeteus sits" – not a lengthy statement, is it?

Theat. No, of very modest length.

Str. Now it is for you to say what is about – to whom it belongs.

Theat. Clearly about me: it belongs to me.

Str. Now take another.

Theat. Namely–?

Str. "Theaeteus (whom I am talking to at this moment) flies".

Theat. That too can only be described as belonging to me and about me.

Str. And moreover we agree that any statement must have a certain character.

Theat. Yes.

Str. Then what sort of character can we assign to each of these?

Theat. One is false, the other true?

Str. And the true states about you things that are (or the fact) as they are.

Theat. Certainly.

Str. Whereas the false statement states about you things *different* from the things that are.

Theat. Yes.

Str. And accordingly states *things that are – not* as being.

Theat. No doubt.

Plato clearly attributes truth and falsehood to sentences as grammatical units having a definite structure: noun + verb. Neither nouns (names) nor verbs (predicates in our current logical sense) can be qualified as true or false. The relevant definitions extracted from both translations are as follows:

(I) A sentence *A* about an object *O* is true if and only if *A* says about *O*, things or facts, that they are, as they are; A sentence *A* about an object *O* is false if and only if *A* says about *O* differently than they are;

(II) A sentence *A* about an object *O* is true if and only if *A* states about *O* things or facts, as they are; A sentence *A* about an object *O* is false if and only if *A* states about *O* differently than they are.

For simplicity, (I) and (II) can be regarded as equivalent under the assumption that the words 'saying' and 'stating' express the same. Conford (pp. 311–314) adds some interpretative comments (p. 310):

> […] the notion [of truth] is that truth consists in the *correspondence* of the statement with the 'things that are' or 'the facts'. How they correspond is not explained.

This explanation looks as made in the name of Plato. The main line is this. Assume that the sentence 'Theaeteus sits' is true. It is so because 'Theaeteus' stands for Theaeteus, 'sits' stands for sitting, and 'Theaeteus sits' corresponds to the obtaining fact, namely: Theaeteus sitting. And we read further (p. 311):

> Here each of the two words *stands for* the one element in the complex fact. The statement as the whole is complex and its structure *corresponds to* the structure of the fact. Truth means this correspondence.

However, Plato says nothing about correspondence and nothing about facts. Thus, Conford's explanation is an obvious over-interpretation of what Plato literally said in *Theaeteus*, very similar to what Michel did with the text from the Bible (see the end of Sect. 1.1). Thus, we should rather deal with Plato's statements about truth that are literally and not ascribe to him views heavily dependent on a truth-definition much later formulated. Conford tries to reconcile Plato with the correspondence theory of truth in its modern version—truth as the correspondence with facts. In particular, as the phrase 'correspondence with facts', indicates, he was influenced by the language employed by British philosophers, discussing the concept of truth at the end of the 19th century and at the beginning of the 20th century (see Chap. 3, Sect. 3.5).

There is one point which I deliberately neglected in my foregoing remarks. It is the problem of how false statements are possible. This issue is usually considered when one speaks about Plato's theory of propositional truth (see, for example, Lorenz, Mittelstrass 1966, and a very extensive treatment in Szaif 1996, Part II). Plato considered these problems in *Theaetetus* 187–200 and *Sophist* 237–246, 259–263. This second dialogue clearly shows that Plato's quoted definitions of truth were embedded into the context of his considerations about the possibility of false sentences. I did not enter into *Theaetetus* because I did not find in this dialogue any definition of truth, although Plato speaks there about truth, falsehood and *episteme*, and discusses relations between these concepts. I deliberately kept some fragments from *Sophist* that are associated with the possibility of false sentences. Consider the following fragments:

> (Ia) Visitor: Whenever there's speech it has to be about something. It's impossible for it not to be about something.
>
> (Ib) Visitor: And the false one says d things different from *those they are*. Visitor: So it says *those that are not*, but that they are.
>
> (IIa) Str[anger]: Whenever there is statement, it must be about something; it cannot be about nothing. Str. Whereas the false statement states about you things *different* from the things that are.
>
> (IIb) Str. And accordingly states *things that are – not* as being.

There is an interesting difference in translation. In (Ia), Visitor/Stranger says that if there is a speech about something, it cannot be not about something, but Visitor/Stranger seems to derive in (IIa) a different conclusion about speech: since speech must be about something, it cannot be about nothing. Let us agree for the sake of argument that if I speak about nothing, I do not speak about Theaetetus. Platon insisted that it is impossible to speak about nothing. He was right, provided that classical logic is applied. Saying something about nothing would have to have the form 'nothing is P', where 'nothing' functions as a term. Since nothing as the denotation of 'nothing' has to be individuated, we have further that there is something which is nothing, what yielding a contradiction. This proves that the sentence ‘nothing is *P*’ has no model, that is, it is not about anything. If we conventionally define that ‘not something (anything) = nothing’, we can say that if it is necessary for speech to be about something, then it is impossible for it to be about nothing. Consider now the crucial sentence ‘Theaetetus flies’ in the circumstances adduced by Plato, that is, in the situation where Theaetetus is sitting. One can say that the sentence ‘Theaetetus flies’ is about nothing. Conford says (p. 313) that, according to this interpretation, our sentence is about a non-existing fact (p. 313). Plato ascribes to Visitor/Stranger the following inference: if a sentence states about an object *O* differently than they are, than it entails that things those that are not, are.

The Visitor/Stranger came from Elea and he exposed Parmenides’ theory. Let us assume that the Parmenidean logic licensed the inference (I do not assert that it is so): if ‘*a* is *P*’ is true and ‘*a* is *Q*’ is false, then the latter is about nothing. Certainly, predicate logic invalidates this reasoning. We have only this: if ‘*a* is *Q*’ is false, then ‘it is not the case that *a* is *Q*’ is true (eventually ‘*a* is non-*Q*’ is true). Since individual constants in predicate logic are not empty (it is the strict rendering of Plato’s claim that speech must be about something), we can perfectly well say that the sentence ‘Theaetetus flies’ is about Theaetetus, but it is false (or states false things about Theaetetus, or states about him things that are not). The problem is how to understand phrases, like ‘things that are’. There is no trouble if we regard them as expressing predications, but difficulties arise when a purely existential meaning is attributed to them. Two additional remarks are in order at this place. First, the translation of Plato by Nicolas White (and by others in Plato’s, *Complete Works*, edited by John Cooper and D. S. Hutchinson) is much more ordinary than Conford’s translation. As far as I know translations of the classics in a manner more consonant with ordinary Greek or Latin has now become the standard. This shift can sometimes be decisive for an interpretation; my guess is that White’s translation does not allow interpreting Plato’s theory of truth as a correspondence theory, at least not without further ado. Second, it has become a very common place to say that Plato’s analysis of false sentences was a reply to the Sophists. As a matter of fact, Protagoras (see Sect. 1.2) maintained that there everything (every opinion) was true. However, he did not argue by logical means but rather by the man-the-mensura principle. I do not deny that Plato opposed the Sophists, but it seems that in this case he refuted Parmenides, although is it is not surprising that the Stranger of Elea spoke on Plato’s behalf. Plato’s argument was very simple: if ‘*a* is *P*’ is true and it implies that ‘*a* is non-*Q*’, then ‘*a* is *Q*’ is false. Therefore, false sentences about something are possible.

As a result of the above analysis, I will not follow Conford’s interpretation of Plato on truth, because I have very grave reservations about overusing contemporary conceptual idioms in explaining views expressed by past philosophers. Of course this statement does not mean that looking at Plato (or other old masters) through contemporary lenses should be neglected. However, two things must be distinguished: first, translating past texts; second, comparing them with what is going on in present discussions. The former job requires literal reproductions of linguistic usages practiced in a given historical epoch; the latter task allows more or less free interpretive steps. I think that if we want to do justice to history, we should stay with the simple interpretation of Plato suggested in **DG9**, which consists in saying that a sentence is true if it says about its object as it is, and false if it say as it is not. It assumes that the phrases 'they are' and 'they are not' are understood as abbreviated predications. In a more contemporary terminology, we can express this idea in the following manner: a sentence '*a* is *P*' is true if and only if it says about things as they are, that is, it says that *a* is *P*; otherwise it is false.

1.4 Aristotle

There is no question that Aristotle was the most influential person in the entire history of aletheiology. Below is a selection of the Stagirite's statements about truth (I will mention various translations in cases in which differences seem significant): E—Aristotle, *Categoriae*, *De Interpretatione*, tr. by E. M. Edgill, in *The Works of Aristotle*, ed. by W. D. Ross. Oxford: Oxford University Press 1928; A—Aristotle's, *De Interpretatione*, tr. by J. L. Acrill. Oxford: Clarendon Press 1963; C—Aristotle, *On Interpretation*, tr. by H. P. Cooke, in *Aristotle* in *Twenty-Three Volumes* I. Cambridge, Mass.: Harvard University Press 1983; R—Aristotle, *Metaphysics*, tr. by W. D. Ross. Oxford: Oxford University Press; K—Aristotle, *Metaphysics*, Books *Γ*, *Δ*, and, *E*, tr. by Ch. Kirwan. Oxford: Clarendon Press; due to Aristotle's significance for the Middle Ages, I also include Latin translations (indicated by (L)) of some fragments after *Aristoteles Latine*, edidit Academia Regia Borussica. Berolina: Georgius Reimer 1831):

(A1) *Categorie* 14b

(E) The fact of the being of a man carries with it the truth of the proposition that he is, and the implication is reciprocal: if a man is, the proposition wherein we allege that he is, is true, then he is. The true proposition is, however, is in no way the cause of the being of the man, but the fact of the man's being does seem somehow to be the cause of the truth of the proposition, for the truth or falsity of the proposition depends of the fact of the man's being or not being.

(A2) *De Interpretatione* 9a30

(E) A sea-fight must either take place to-morrow or not, but it is not necessary that it should take place to-morrow, neither is necessary that it should not take place, yet it necessary that it either should or should not take place to-morrow. Since propositions correspond with facts, it is evident that when in future events there is a real alternative, and a potentiality in contrary directions, the corresponding affirmation and denial have the same character.

(A) (Omitting the first sentence) […], since statements are true according to how the actual things are, it is clear that wherever these are such as to allow of contraries as chance has it, the same necessarily holds for contradictories also.

(C) (Only the middle sentence) […] as the truth of propositions consists in corresponding with facts.

(L) (Only the middle sentence) […] cum orationes similiter verae sint atque res [….].” (tr. by Julius Pacius)

(A3) *Metaphysics* 1011b

(R) Falsehood is saying of that which is that it is not, or that what which is not, that it is; truth is saying of that which is that it is, or of that which is not that it is not. Therefore he, who says that a thing is or not is, says what is either true or false. But if the subject is a middle term between contradictories, neither that which is nor which is not is being said to be or not to be.

(K) […] for to say that that which is not or that which is not is, is a falsehood; and to say that that which is, is and that which is not, is not, is true; so that, also, he who says that a thing is or not will have the truth or be in error. But it is said that neither that which is nor that which is not, either is not or is.

(L) […] *dicere namque ens non esse aut hoc esse, falsum: ens autem esse et non esse, verum est. Quare et qui dicit esse aut non esse, verum dicet aut mentietur, sed nec ens dicitur non esse aut esse, nec non est.* (tr. by Bessarion)

(A4) *Methaphysics* 1027b

(R) But since that which *is* in the sense of being true, or *is not* in the sense of being false, depends on combination and separation, and truth and falsity depend on allocation of a pair of contradictory judgments (for the true judgement affirms when the subject and predicate really are combined, and denies when they are separated, while the false does the opposite of this allocation; it is another question, how it happens that we think things together or apart […]. […] falsity and truth are not in things […], but in thought.

(K) That which is as true and that which is not as falsehood are concerned with composition and division and, taken together, with the apportionment of a contradiction. For truth has the affirmation in the case what is compounded and the denial in the case of what is divided, while a falsehood has the contradictory of this apportionment. (How we come to conceive things together or separately is another question) […]. […] Falsehood and truth are not in actual things […] but in thoughts.

(L) […] *verum, ens et non ens, ut falsum quonian circa compositionem et divisionem est, et omnino circa partitionem contradictionis. Verum etenim affirmationem in compositio habet, negationem vero in diviso, falsum vero huius partitionis contradictionem quo autem modo quod simul aut quod separatim est, intelligere accidat, alia ratio est.* […] *Non enim est est falsum et verum in rebus* […] *sed in mente.* (tr. by Bessarion)

(A5) *Metaphysics* 1051 a–b

(R) ‘Being’ and ‘not being’ are employed firstly to with reference to the categories, and secondly with reference to the potency and actuality […], and thirdly in the sense of true and false. […] he who thinks, This depends, on the side of objects, on their being separated and the being combined, so that he who thinks the separated to be separated and the combined to be combined has the truth, while he whose thought is a state contrary to that of the object is in error. […]. It is not because we think truly that you are pale, that you *are* pale, but because you are pale we who say this have the truth.

(L) (Restricting to the fragment on combined and separated) *Quamombren verum dicit, qui divisum dividi et compositum componi putat, falsum autem, qui contraquam res se habeant, aut quando sunt aut not sunt*. (tr. by Bessarion)

Aristotle clearly distinguishes being *qua* being and being *qua* truth (see (A5)). Since truth and falsehood is in thoughts (see (A4)), not in things, we can interpret this distinction as one between ontological and epistemological concepts of truth, and regard Aristotle as advocating the epistemological account of truth. Another doubtless point is that the Stagirite closely connects truth, falsehood and logic, defending (see (A3)) the principles of contradiction and of the excluded middle, that is, bivalence (see, however, (**DG17**) below). Truth is independent of concrete acts of thinking (see (A1), (A5)) which do not cause that something is true or false; the causal relation, as commonly formulated, goes rather from world to truth. Aristotle also (see (A4)) makes a distinction between truth and its criteria. Although he is not quite explicit about bearers of truth, we can assume that sentences, judgments, propositions or statements play this role (see (A2); in the Greek original the word *logoi* occurs). In many respects Aristotle continues Plato's path in thinking about truth and fortifies; he also adds something new, particularly as far as logic is concerned. After all, his theses about this topic are ordered and interconnected, and this decides that we are here to do dealing the first full-blooded theory of truth in the history of philosophy.

Fragments of (A3), (A4) and (A5), more precisely the statements (I repeat) (a) Falsehood is saying of that which is that it is not, or that what which is not, that it is; truth is saying of that which is that it is, or of that which is not that it is not; (b) True judgement affirms when the subject and predicate are in fact combined, denies when they are separated, while the false does the opposite; (c) Truth means thinking that to be divided or united which is divided or united, respectively; error means being in state contrary to the facts, look like attempts to define the concept of truth. The statements (b) and (c) express approximately the same. However, given a literal interpretation, (a) is different. The main point is that (a) applies to existential propositions (*a* exists, or *a* does not exist), but (b) and (c) refer to the composition of subject and predicate ('*a* is *b*', '*a* is not *b*'). The interpretation of (b) and (c) is much easier. Let us say that the structure '*a* is *b*' express that the subject *a* and the predicate *b* are combined, and the structure '*a* is not *b*' that they are divided. It is the level of language. On the level of things, we have combination and division of substances, the primary one denoted by *a* and the secondary—by *b*. Now, we can say that a sentence is true if combination (division) on the level of language is properly related to the combination on the level of substances; otherwise, it is false. This explanation cannot be extended to existential sentences in a straightforward way. Why did Aristotle introduce this dualism? One explanation is that (a) appears in the context of Aristotle's polemic against Protagoras. The formula (a) is an extension of the man-the-mensura principle (see Kahn 1973, p. 367 to the effect that Protagoras introduced the prototype of (a) into Greek philosophy), and Aristotle simply used the language of his philosophical adversary when he criticized him. Aristotle's task consisted in defending the principle of the excluded middle against sophistic attacks, and perhaps he wanted to execute the defence by adopting Protagoras' truth-talk. However, if we accept Kahn's view (see Kahn 1973, pp. 331–370) about the veridical use of the Greek counterpart of 'be' (*esti*, *einai*), then a new perspective opens up. According to Kahn (see p. 367) the formula "that what (instead "which is", but it does not make any difference) is" is to be understood as "what is so". With this hint, (a) means (see Kahn 1973, p. 336): (a_1) To say of what is (so) that it is and of what is (not so) that it is, is falsehood; to say of what is (so) that it is and of what is not (so) that it is not, is truth. This reading makes (a) closer to (b) and (c), and we can now say that the latter explain the meaning of 'what is (is not) so'.

A separate, but related problem concerns whether Aristotle offered a version of the correspondence theory of truth (a positive answer is defended Crivelli 2004, Chap. IV. According to Cavini 1993, p. 87, (A1) is an informal statement of the equivalence of the type '*A* is true if and only if *A*' (**T**-equivalence; see Chaps. 7 and 8). This point is important because **T**-equivalences (or **T**-biconditionals or **T**-sentences) are a stable ingredient of any correspondence theory of truth. However, two remarks are in order here. First, as I will later argue in the book, **T**-equivalence is only a part of Tarski's theory, not its whole. Second, Cavini's interpretation must be qualified by observing that a causal nexus is involved in the Aristotelian formulation, but later (contemporary) versions of **T**-equivalences are devoid of any such factors. On the other hand, the idea that the truth of a sentence is rooted in the world, but facts are independent of their assertions, was always accepted by advocates of this theory. (A2) is of special interest in this context. The relevant fragment goes in Greek original (Latin transcription) reads: (G) *est epei honoios oi logoi aleteis hosper ta pragmata*. It is the first part of a compound sentence the second part of which asserts that both contrary contingent statements about the future (that is, 'sea-battle will be tomorrow' and 'sea-battle will not be tomorrow') have the same character, that is, are possible (I will not get into the famous problem of future contingents). The fragment (E) translates (G) as "Since propositions correspond with facts", which is obviously wrong for neglecting the word *aletheis*; clearly, (G) says something about true propositions, not propositions. The problem with (A) and (C) is that the connection between both parts of the whole sentence is mysterious, for it is not clear how truth as the correspondence with facts, or consisting of how the actual things are altogether makes future contingents possible. Independently of that, one may note that no word in (G) justifies using of the word 'correspondence', 'corresponds' or *similariter* (as in (L)). In this respect, (A) is much better, but it suggests too much, because it employs the phrase 'the actual things'—which is burdened by quite rich philosophical contents. I think that should be translated literally, something like 'just because statements are true according to how things (events, facts, etc.) are'. It clearly indicates that the double possibility of contrary future contingents depends of how things, are or rather will be in the future. Nothing more needs to be added to explain the situation.

J. L. Ackrill says in his commentaries to (A): "[Aristotle] seems to hold a rather crude realistic correspondence theory of truth" (p. 140). Ackrill (as many other commentators; I will mention them later, but see Carretero 1983 for an extensive treatment of Aristotelian theory as based on the concept of correspondence) takes this qualification for granted as beyond all doubt. However, I see little justification for this interpretation. Certainly (G) is not sufficient to justify Acckrill's claim.

More promising is (A4) and (A5). Both seem to invite the concept of correspondence, but as an interpretative device, and not as something provided by the literal meaning of those fragments. A good summary of the problem is to be found in Kahn 1973, p. 367:

> As we have seen […] the classical formula given by Aristotle – to say of what it is and of what is not that is not – merely articulates the pattern of the ordinary veridical idiom in Greek. Wherever their full structure is clear, these uses of *εἰμί* are characterized by an explicit comparison, formulated by *οὕτως … ως* between an essive clause which expresses how things are, were, or will be, and an intentional clause with a verb of saying or thinking. […] As in the most contemporary idiom so in Homer and Sophocles: the man who speaks the truth "tells it like it is", and the liar tells it otherwise.

This informal *façon de parler* leaves open many issues involved in a correspondence theory which conceives truth as relating language to the world. This idiom only specifies that there is a relation of this kind, and that it admits a comparison between its terms (relata). In particular, one term is to be found in *what is said or thought*, but the other one in *what is actually the case*, and that the truth depends upon some point of similarity or agreement (*οὕτως … ως*) between the two.

Yet I think that to use words 'agreement', 'similarity' and 'comparison' suggests too much, because they are dangerous words in explaining philosophical issues related to the concept of truth, unless they are understood in a very informal sense. But if it is the case, I would cancel the word 'comparison' and the very last clause, namely, 'that truth depends, etc.' The formula (a_1) expresses the basic Aristotelian intuition concerning the concept of truth, which, as the subsequent history shows, was considerably obscured by explanations like (A4) and (A5). It interesting that (a_1) can be seen as a development of the classical scheme: *verbum dicendi* plus *aletheia*. In fact, the formula

> To say of what is not (so) that it is and of what is that it is not (so), is falsehood; to say of what is (so) that it is and of what is not (so) that it is not, is the truth,

seems to explain the meaning of *aletheia* hidden in traditional usages, ordinary as well as philosophical. In order to see this link it is enough to rephrase (a_1) as "To say falsehood is to say what of is (so) that it is and of what is not (so) that it is; to say truth is to say of what is (so) that it is and of what is not (so) that it is not." Aristotle's step meant the last step toward an abstract treatment of *aletheia*.

(**DG17**) The translation of (G) shows an explicit tension between philosophical faithfulness of translation and its dependence on philosophical ideas used by interpreters, consciously or not. It is obvious that translations of (G) as in (E), (C) and (L) (less in (A)) are strongly related to the tradition that views the Aristotelian account of truth as the full correspondence theory. I decisively prefer more literal translations of classical philosophical texts over those of their interpretations that are guided by later or contemporary (to the translators) philosophical theories.►

(**DG18**) I omitted many important points in Aristotle. For example, some authors (notably Jan Łukasiewicz) doubted whether Aristotle defended the principle of the excluded middle in its full scope. That concerns of course the problem of future contingents which I deliberately skipped (see Chap. 9, Sect. 9.5 for a treatment from the point of view of **STT**). However, the matter is not without significance for general aspects of the Aristotelian truth theory. We have pieces of evidence that the Stagirite accepted an absolutist perspective on truth: truth-values are time-independent. On the other hand, if future contingents are neither true nor false at the present time, the situation becomes more complex, because we need to decide whether such assertions do or do not have logical values. Additional points that deserve attention are: the ontological concept of truth and truth as related to non-composites. However, I was only interested in extracting the core of Aristotle's theory of truth regarding points that are relevant for further discussions—or better, are more relevant than others. Moreover, I will come back to some neglected issues in the systematic part of this book. I will also clarify the problem of the extent to which Aristotle formulated the correspondence theory of truth via the distinction between two concepts of correspondence (see Chap. 9, 9.2). We also have (see Long 2011) an ontological interpretation of Aristotle's theory of truth, which is, largely based on Heidegger and late phenomenology. For reasons explained in **DG3**, I entirely reject such a reading of the Stagirite as a complete arbitrary interpretation based on unjustified projections of very dubious philosophical presuppositions.►

1.5 Ancient Philosophy After Aristotle

The most important ancient semantic theory after Aristotle was produced by the Stoics, particularly by Zeno of Kition and Chrysippus (the former continued ideas of the Megarian School. Unfortunately, knowledge of it is fragmentary and second-hand. According to Sextus Empiricus (*Against the Mathematicians* 8.70; quoted after Barnes 1993, p. 55):

> The Stoics claimed […] that the true and the false are found in sayables. And they say that a sayable is that which subsists in accordance with a rational presentation [that is] can be set out in language.

Although there is a controversy (see Barnes 1993, Barnes 2007, *passim*) over how to interpret the concept of sayables—as a mental entity or as an objective item (an abstract content expressed in words—it is rather indisputable that sayables, as bearers of truth, belong to propositional (or sentential) category. The Stoics radically defended bivalence, that is, the thesis that every truth-bearer is either true or false (see Barnes 2007, Chap. 1). Several references to the Stoics in secondary sources (Cicero, Diogenes Laertius, Sextus Empiricus, Galen, Alexander of Aphrodisias) collected in Cavini 1993 lead to the following formulation (see Cavini 1993, p. 93; he says it “expresses the Stoic ‘semantic’ definition of truth or Correspondence Thesis”):

(S) If x says that p and p, then x is speaking truly.

There is an obvious link between (S) and the tradition. First of all, (S) employs the traditional scheme *verbum dicendi* plus *aletheia*. Second, (S) can be viewed as a version of Aristotle’s (A3). Third, (S), even if regarded as a Correspondence Thesis, makes no appeal to ontological constraints, like unity, diversity, etc. Thus, (S) expresses a purely semantic notion of truth—truth is a property of utterances.

(**DG19**) Cavini 1993, p. 94 argues that a formula given by Alexander of Aphrodisias, namely *o to on einai legon alethei* (Who says that it is, when it is, speaks truly), may be attributed to the Stoics. Although this sentence does not contain variables, it is rather clear that the phrase ‘it is’ functions as a schematic letter (in the sense of contemporary logic) and that fact justifies a semiformal formulation of (S) with usage of variables. From the contemporary point of view, (S) anticipates several important ideas. In particular, we can reverse it (with a small addition) in the form:

(S1) If x is speaking truly that p, then x says that p and p.

Now, combining (S) and (S1) leads to:

(S2) x says that p and p if and only if x speaks truly that p.

At the first sight, (S2) is similar to Tarski’s **T**-scheme, that is, the formula ‘*A* is true if and only if *A*’ (it is a preliminary formulation; see **DGIII6** and Chap. 7, Sect. 7.4 for further explanations. In fact, (S2) better anticipates the minimalist theory of truth (see more about this account in Chap. 3, Sect. 3.6). (S) is much closer to **T**-scheme (in Tarski’s sense) than (A1) because it does not involve any causal relation between acts of thinking and truth (see Sect. 1.4 above), and can actually be interpreted as a prototype of **T**-equivalence; comparing (A1) and (S) shows that the level of abstraction in the Stoics was much higher than in Aristotle. In particular, Stoic logic was more sophisticated than that of the Stagirite. In particular, the Stoics developed a considerable part of propositional calculus as a theory more fundamental than the syllogistic. Unfortunately, the general neglecting and misunderstanding of Stoic logical and semantic ideas until the twentieth century prevented their wider influence, and probably considerably stifled the development of logic and semantics.►

(**DG20**) The Epicureans developed another approach to truth. According to them, sensations are proper bearers of truth. It was probably a return to an older tradition, for example, which represented to some extent by Democritus, although he considered the senses rather as a source of error than correct cognitive results. A novelty of Epicurean philosophy consisted in the view that evidence, provided by senses, was a truth-criterion. The Sceptics, as we see from the works of Sextus Empiricus, were very much interested in truth. Although their interest was mainly negative, serving their criticism of various dogmatisms in philosophy, they preserved a great deal of important data about the views of other philosophers; in fact, the works of Sextus are one of the most important sources for our knowledge of ancient philosophy. As is well-known, the Sceptics criticized any possibility of a reliable truth-criterion. Hence, their considerations became important for any attempt at formulating such a criterion. Carneades, as reported in Sextus Empiricus, *Against the Logicians* I 168–169, spoke about true and false presentations. According to Carneades, every presentation points out an aspect of a presented object. A presentation is true if it agrees with what is presented, and false otherwise. Of course, this formulation of what truth is, anticipates the later, explicit correspondence jargon. Carneades himself belonged to the Academicians, a philosophical camp close to scepticism. Perhaps it is also important to note that the Sceptics' criticism of truth-criteria was addressed to views that assumed the concept of *episteme* as absolute knowledge.

Aurelius Augustinus is the last philosopher to be treated in this section. He developed a strong ontological view concerning truth (see Hessen 1931, pp. 27–60 for an extensive treatment of Augustinus' epistemology, in particular, his truth-theory, and its connections with metaphysical issues. In general, Augustinus managed successfully the first synthesis of theology and philosophy that was acceptable to the Catholic Church; since almost one Augustinus' views has two dimensions: theological and philosophical. The same applies to his conception of truth, and, in order to understand them, one should try to separate out the theological and philosophical factors, although both are closely interconnected. The theological dimension of Augustinus' theory of truth is best represented by the following passage ((*De doctrina Christiana* 8, 115–120):

> *Quamquam nemo debet aliquid sic habere quasi suum proprium, nisi forte mendacium. Nam omne verum ab illo est, qui ait: Ergo sum veritas* (Nobody should consider anything as his own, perhaps with exception of lying. All truth comes from Him who said: I am truth.)

This view has far-reaching consequences. First, knowledge of truth is God-oriented, that is, the ultimate reason to search for truth consists in the attempt to achieve knowledge of God. On the other hand, the famous doctrine of *illuminatio* as a necessary condition of knowledge (understood as *episteme*) claims that truth is not graspable without God's help, executed by acts of His free decision (grace). Consequently, this account of truth and its accessibility is more theological than philosophical.

However, Augustinus also developed a more secular account of truth. It is expressed in *Soliloquia* II, 8:

> (a) *Verum est quod ita habet ut cognitor videtur, si velit possitque cognoscer* (Every thing is true which is thought by someone who wants and can know it);
>
> (b) *Verum mihi videtur esse id, quod est* (I think that everything is true, that is).

At the first glance, the formulation (b) is a concise summary of (a). That is true, but only to an extent. Let me start with (b). It is clear that it offers a fairly ontological account of truth because we have here the equality: (c) a is true = a is (exists), and since the letter *a* in (c) ranges over things, we cannot apply the veridical use of 'is' to the interpretation of (b). Thus, the epistemological reading of (b) is ruled out. In this respect, Augustinus returned, probably unconsciously, or perhaps via Plato, to Parmenides' identity of truth and being. Anyway, Augustinus certainly revived the ontological theory of truth in the spite of the dominance of epistemological accounts in the post-Parmenidean era.

As long as we focus on the first words of (a), that is—*Verum est quod ita habet* —we are on the purely ontological level, but the second part—namely *ut cognitor videtur*, *si velit possitque cogonscer*—seems to add something new. An epistemological interpretation is tempting, and of course possible. However, I do not think that (a) can be interpreted as expressing a view akin to that of Aristotle's, for example. It is rather the case that Augustinus states that idea of knowledge which was related to his view about *illuminatio* and God's grace as necessary conditions for a successful searching for truth. It is why we cannot ignore the theological dimension in Augistunus theory of truth. The same comment applies to another interesting passage (*Soliliquia* II, 2), namely—*Erit igitur veritas*, *etiamsi mundus intereat*—which says that truth is timeless, unchangeable and unconditional. Augustinus does not derive these properties of truth via epistemological or ontological considerations, but rather by appealing to the absolute stability of God. In fact, if all truth comes from God and He is truth, the mentioned properties are obvious and do not need be separately proven. These linkages between theology and theory of truth, although not relevant to the ideas developed by Plato, Aristotle or the Stoics, became important in the Middle Ages for their contribution to making coherent (in a sense, as we will see) the theory of transcendentals, in particular the relation between Being and Truth.

(**DG21**) One of the most famous passages about truth occurs in the *New Testament* (John 37–38). The text reads as follows:

> "You are the king, then!" said Pilate. Jesus answered, "You say that I am a king. In fact the reason I was born and came into the world to testify to the truth. Everyone on the side of truth listens to me."
>
> "What is truth?" retorted Pilate. With this he went out again to the Jews gathered there and said, "I find no basis for a charge against him."

Many contemporary philosophers use Pilate's question as the motto in writings that discuss the problem of truth in general. However, I have some doubts whether Pilate was actually interested in the concept of truth as such. I am inclined to the following interpretation of the quoted fragment of the most famous Roman procurator. Pilate was a well-educated Roman lawyer, probably with a fairly good knowledge of philosophy of his time. One can even conjecture that Pilate was influenced by the Sceptics, very popular among Roman intellectuals and officials at the break of Millennia. If that was the case, he must have been surprised by Jesus' reply "I came into the world to testify to the truth", because that went directly against the sceptical rejection of any truth-criteria. Since this philosophical position consisted in abstaining from definite assertions, he did not formulate any charge against Jesus and left the issue to be resolved by the Jews. Of course, this interpretation is only a considerable piece of speculation, but consistent with considering Pilate as a representative of scepticism. If so, his retort was rhetorical, because he had no interests in deliberations about the nature of truth.►

(**DG22**) Legal thought had a considerable impact on the development of philosophy, in particular of philosophical terminology. The case of Roman law is extensively investigated in Giaro 2007. The author shows how legal practices contributed to the development of truth-criteria.►

1.6 Concluding Remarks

Aristotle's theory of truth can be considered as the turning point in the history of aletheiology in at least two respects. First, his ideas were presented in an abstract way, contrary to his predecessors who used a more or less figurative language (*Metaphysics* can be regarded as the first academic textbook of philosophy—its style is very professional, even from the contemporary point of view. Second, Aristotle's theory of truth had explicit links with logic. This fashion of doing truth-theory became even more explicit in the case of the Stoics. Specific contributions of ancient aletheiology, important from the point of view of **STT** (and also that from other philosophical theories of truth), can be summarized in the following points:

1. The idea that truth consists in saying how thing are (many authors from various cultures, philosophical, religious, literary, etc.);
2. The problem of truth-criteria (the Sceptics);
3. Is truth absolute or relative? (Protagora and his critics, particularly Plato);
4. The principle of bivalence (Aristotle);
5. Truth and being (Parmenides, Plato);
6. The idea that truth is in thought not in things (Aristotle);
7. Truth and knowledge (*episteme*, *doxa*, Parmenides, Plato);
8. Truth and morality (various answers to the question what is moral truth and how could be recognized (all ancient philosophers).

References

Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E., Romanowska, A., Woleński, J. (Eds.) (2004). *Provinces of Logic Determined. Essays in the Memory of Alfred Tarski. Parts I, II and III* (*Annals of Pure and Applied Logic*, 127). Amsterdam: Elsevier.

Assmann, J. (1990). *Ma'aat. Gerechtigkeit und Unsterblichkeit in Alten Ägypten*. München: C.H. Beck.

Barnes, J. (1993). Meaning, saying and thinking. In Döring, Ebert (1993), 47–61.
Barnes, J. (2007). *Truth, etc. Six Lectures on Ancient Logic*. Oxford: Clarendon Press.
Bello, A. G. A. (2004). Some methodological controversies in African philosophy. In Wiredu (2004), 263–273.
Boeder, H. (1959). Der frügrechische Wortgebrauch von Logos und Aletheia. Archiv für Begriffsgeschichte, 4, 82–112.
Burnet, J. (1892). *Early Greek Philosophy*. London: Adam and Charles Black.
Bultmann, R. (1933). Der griechische und hellenistische Sprachgebrauch von άκήheia. In Kittel (1933), 239–240.
Carretero, F. M. L. (1983). *Der aristotelische Wahrheitsbegriff und die Aufgabe der Semantik*. Köln: Universität zu Köln.
Cavini, W. (1993). Chrysippus on speaking truly and the Liar. In Döring, Ebert (1993), 85–10.
Classen, C. J. (1989). Protagoras' *aletheia*. In Huby, Neal (1989), 13–38.
Conford, F. M. (1935). *Plato's Theory of Knowledge. The Theaetetus and the Sophist of Plato with a Running Commentary*. London: Kegan Paul, Trench, Trubner.
Conford, F. M. (1952). *Principium Sapientiae: The Origin of Greek Philosophical Thought*. Cambridge: Cambridge University Press.
Crivelli, P. (2004). *Aristotle on Truth*. Cambridge: Cambridge University Press.
Döring, K., Ebert, T. (Eds.) (1993). *Dialektiker und Stoiker. Zur Logik der Stoa und ihrer Vorläufer*. Stuttgart: Franz Steiner Verlag.
Detienne, M. (1967). *Les Maîtres de vérité dans la grèce archaïque*. Paris: François Mason; Eng. tr. (by J. Lloyd), *The masters of truth in Archaic Greece*. Cambridge, MA: The MIT Press 1996.
Friedländer, P. (1954). *Platon: Seinswahrheit und Lebenswirklichkeit*. Berlin: de Gruyter; Eng. tr. (by H. Meyerhoff), Plato, v. 1-3. London: Routledge and Kegan Paul.
Frisk, H. (1936). *"Wahrheit" und "Lüge" in den indogermanischewn Sprachen. Eine morphologische Betrachtungen*. Göteborg: Elanders Boktryckeri Aktiebolag.
Giaro, T. (2007). *Römische Rechtswahrheiten. Ein Gedankenexperiment*. Frankfurt am Main: Vittorio Klosterman.
Guthrie, W. K. C. (1969). *A History of Greek Philosophy*, v. 3: *The Fifth-Century Enlightenment.* Cambridge: Cambridge University Press.
Hallen, B. (2004). Contemporary Anglophone African philosophy: A survey. In Wiredu (2004), 99–148.
Heidegger, M. (1940). Platons Lehre von der Wahrheit. *Geistige Überlieferung*, 2, 96–124.
Heitsch, E. (1962). Die nicht-philosophische ΑΛΗΘΕΙΑ. *Hermes. Zeitschrift für klassische Philologie*, 90, 24–33.
Herbertz, R. (1913). *Das Wahrheitsproblem in der giechischen Philosophie*. Berlin: Reimer.
Hessen, J. (1931). *Augustinus Metaphysik der Erkenntnis*. Berlin: Ferdinand Dümmler Verlag.
Huby, P., Neal, G. (Eds.) (1989). *The Criterion of Truth*. Liverpool: Liverpool University Press.
Kahn, C. (1973). *The Verb 'Be' in Ancient Greek*. Dordrecht: Reidel.
Kittel, G. (Ed.) (1933). *Theologisches Wörterbuch zum Neuen Testament*. Stuttgart: Kohlhammer Verlag.
Komornicka, A. (1979). *Studia nad Pindarem and archaiczną liryką grecką. W kręgu pojęć prawdy i fałszu* (Studies on Pindar and Archaic Greek Lyric. Around the Concepts of Truth and Falsity). Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
Krischer, T. (1965). ΕΤΥΜΟΣ und ΑΛΗΘΝΣ. *Philologus*, 109, 161–174.
Kumari, K. (1987). *Notion of Truth in Buddhism and Pragmatism*. New Delhi: Capital Publishing House.
Levet, J. P. (1976). *Le vrai et le faux dans la pensée grecque archaïque. Étude de vocabuilare*. Tome 1: *Présentation générale. Le vrai et le faux dans les épopées homériques*. Paris: Les Belles Letters.
Long, C. R. (2011). *Aristotle on the Nature of Truth*. Cambridge: Cambridge University Press.

Lorenz, K., Mittlestrass, J. (1966). Theaitetos fliegt. Zur Theorie wahrer und falscher Sätze bei Platon (Soph. 251d–263d). *Archiv für Geschichte der Philosophie*, 48, 113–152.
Luther, W. (1935). *"Wahrheit" und "Lüge" in ältesten Griechentum*. Borna: Robert Noske Verlag.
Luther, W. (1965). Wahrheit, Licht, Sehen und Erkennen im Sonnengleichnis von Platonis Politeia. Ein Auschnitt aus der Lichtenmetaphysik der Griechen. *Studium Generale*, 18, 479–496.
Luther, W. (1966). Wahrheit, Licht und Erkenntnis in der griechische Philosophie bis zu Demokrit. *Archiv für Begriffsgeschichte*, 10, 2–270.
McLeod, A. (2016). *Theories of Truth in Chinese Philosophy: A Comparative Perspective*. Lanham: Rowman and Littlefield.
Michel, D. (1968). Ämät. Untersuchungen über "Wahrheit" in Hebräischen". *Archiv für Begriffsgeschichte*, 12, 30–57.
Pelettier, F. J. (1990). *Parmenides, Plato and the Semantics of Not-Being*. Chicago: The University of Chicago Press.
Perzanowski, J. (1996). The way of truth. In Poli, Simons (1996), 61–130.
Peters, F. E. (1967). *Greek Philosophical Terms. A Historical Lexicon*. New York: New York University Press.
Poli, R., Simons, P. (Eds.). (1996). *Formal Ontology*. Dordrecht: Kluwer.
Quell, G. (1933). Der at.liche Begriff "ämät". In Kittel (1933), 233–237.
Rüstow, A. (1910). *Der Lügner. Theorie, Geschichte und Auflösumg*. Leipzig: Teubner Verlag.
Snell. B. (1975). ΑΛΗΘΕΙΑ. *Würzburger Zeitschsrift für die Altertumwissenschaft. Neue Folge*, 1, 9–17.
Szaif, J. (1996). *Platos Begriff der Wahrheit*. Freiburg: Karl Alber Verlag.
Van De Mieroop, M. (2015). *Philosophy before the Greeks: The Pursuit of Truth in Ancient Babylonia*. Hoboken, NJ: Wiley-Blackwell.
Vernant, J.-P. (1983). *Myth and Thought among Greeks*. London: Methuen.
Vroom, H. M. (1989). *Religions and Truth: Philosophical Reflections*. Amsterdam: Rodopi.
Wiesner, J. (1996). *Parmenides. Der Beginn der Aletheia. Untersuchungen zu B2–B3–B6*. Berlin: de Gryuter.
Wiredu, K. (Ed.) (2004). A Companion to African Philosophy. Oxford: Blackwell.
Woleński, J. (1994a). Contributions to the History of the Classical Truth-Definition. In Prawitz, Skyrms, Westerståhl (1995), 481–495; repr. in J. Woleński (1999), 139–149.
Woleński, J. (2004f). *Aletheia* in Greek Thought until Aristotle. In Adamowicz, Artemov, Niwiński, Orłowska, Romanowska, Woleński (2004), 339–360.
Wójcik, A. S. (2010). *"Prawda" w Biblii Hebrajskiej. Analiza logiczno-lingwistyczna* (Truth in Hebrew Bible. An Logico-Linguistic Analysis). Kraków: Austeria.

Chapter 2
Truth from Anselm of Canterbury to Kant

Abstract This chapter continues historical investigations covering the period from Anselm of Canterbury to Kant, including (among others) Abelard, St. Tomas Aquinas, Gassendi, Descartes, Spinoza, Locke, Leibniz, and Hume. The special attention is paid to the history of the formula *veritas est adequatio intellectus et rei.*

2.1 Anselm of Canterbury

Although Arab philosophers, particularly Avicenna (see Sect. 2.4 for some remarks), said some interesting things about truth in the early Middle Ages, I immediately Anselm of Canterbury as a very important truth-philosopher for next generations. Thus, it is justified to begin the history of the concept of truth in medieval philosophy with the latter. Anselm introduces two notions of truth (see Enders 1999 for an extensive discussion and analysis of Anselm's philosophy of truth). First, he says that God is truth, which for this reason is eternal. It is essentially a continuation of Augustinus' theological way of thinking about truth; I will leave this understanding without further comments. Second, Anselm predicates truth about sentences (*orationes*, *enuntiationes*) that express (*significant*) something; we can say that propositions are *significationes* of sentences. Anselm speaks about truth of sentences in two ways. One is represented by the following fragment (*De veritate* 2):

> M. Quando est enuntiatio vera? D. Quando est quod enuntiat, sive affirmando sive negando.
>
> M. Quid igitur tibi videtur ibi [in oratione] veritas? D. Nihil aliud scio nisi quia cum significat esse quod est, tunc est in ea veritas et est vera.

These two exchanges are separated by an important statement:

> Quia si hoc esset, semper esset vera, quoniam eadem manent omnia quae sunt in enuntiationis definitione, et cum quod enuntiat, et cum non est. Eadem enim est oration et eadem significatio et cetera similiter.

Thus, sentences express what there is or is not. Now, a sentence is true if its *significatio* is identical with what is or is not. This is clearly an ontological account

J. Woleński, *Semantics and Truth*, Logic, Epistemology, and the Unity of Science 45, https://doi.org/10.1007/978-3-030-24536-8_2

of truth that anticipates contemporary identity theories (see Chap. 3, Sect. 3.8), and can be justly regarded as a form of the correspondence account of truth.

Anselm also refers to truth of sentences as their *rectitudo* (rightness) (*De veritate* I2): *Ergo non est illi aliud veritas quam rectitudo.* In fact, *rectitudo* is a mixed concept with epistemic and deontic-doxastic dimensions. This just means that *rectitudo* is an evaluative category, characteristic, using contemporary terminology, for the ethics of beliefs. At any rate, according to Anselm, a sentence is true in the ontological sense if and only if it is true via the *rectitudo* criterion.

2.2 Abelard

In Abelard's *Logica Ingredientibus* (290–291) we have the following passage:

> (a) […] vere quaendam convertuntur invicem tamquam causae et effectus, ut essentia hominis et veritas propositionis quae hominem esse enuntiat, hocest '*esse hominem*' convertitur ad orationem veram de se secundum consequentiam essentiae, idest commitationem alternae permanentie, eo videlicet quod oratione hic '*homo est*' proponente hominem esse non potest ipsa in eo esse vera quid homo sit, commitationem itaque Aristoteles accepit inter veritatem propositionis et eventum rei.

Abelard introduces here the subtle problem of how truth is related to what there is. In De Rijk 1956a, p. LII we find the following reconstructs Abelard's view in this way:

> if a proposition is true, then the state of affairs (*eventus rei*) referred to by it exists, and if the state of affairs referred to by a proposition exists, is true.

Now we have the problem of whether both these implications can be accepted together, that is, whether the following equivalence is valid

> (**Ab**) *A* is true if and only if the state of affairs expressed by *A* exists.

The first implication is usually accepted as obvious, but there is a problem with the second. Some medieval logicians introduced the additional condition (called *constantia*) that the proposition in question must be pronounced. However, (in *Logical Ingredientibus*) Abelard had doubts also about the first implication, and proposed to add the *constantia* in this case too. He returned to this problem in *Dialectica* 371, and said:

> (b) […] si ita est in re ut dicit propositio, tunc vera est ipsa propositio,

as saving the equivalence (**Ab**)—provided that the *constantia* is present. Under this proviso we have (De Rijk 1956a, p. LIV):

> (**Ab1**) *A* is equivalent to '*A* is true' if and only if the state of affairs expressed by *A* exists.

Why is the *constantia*s important? It is connected with Aristotle's observation that if it is not so that the sentence that so and so is a cause of what is so and so, but the reverse implication holds. Abelard illustrates the problem by the following example:

(c) […] si ‘Socrates est homo’ dicit illud quod est <in re> et illud quod ipsa dicit sit tantum ‘Socrates est homo’, tunc Socrates est homo.

Videtur tamen posse probari proposita consequentia, haces scilicet:

‘si ‘Soctrates est homo’ dicit illud quod in re est, tunc illud quod in re est dicitur ab ipsa’;

quare et illud quod dicitur ab ipsa in re est; unde ‘*Socrates est homo*’ in re est, quippe, is tantum ab ipsa dicitur. Si vero in re est ‘Isocrates est homo’ ipse re vera homo est.

We can extract the following instantiation of (**Ab**):

si ‘Socrates est homo’ dicit illud quod est <in re> , propositio ‘Socrates est homo’ vera est.

There is a delicate problem with translating Abelard. The formula (b) says:

(*) If it is such (or: if the thing is such) it is said in a sentence, then the sentence in question is true.

If the *constantia* is added we can strengthen the last formula to the equivalence ‘a sentence is true if and only if things are as it says’. The part <*in re*> in (*) is clearly an interpolation by the editor. Is it correct? The literal translation without the interpolated phrase runs as follows:

(**) ‘Socrates is a man’ is true if and only if it says how things are (namely that Socrates is a man).

If this translation is correct it shows that the interpolation by the editor obscures Abelard’s view. The same should be said about (c), which is too philosophical in the sense that it ascribes to Abelard much later views. We can say that (c) is a definition of truth and (*)—its particular instantiation. Abelard’s view is extremely interesting for at least two reasons. First, he observes that the definition of truth applies to real, that is, pronounced sentences. It is an anticipation of the view that truth is relativised to a language. Second, Abelard seems to distinguish a truth-definition requiring a technical vocabulary (this role plays the expression *in re*) and its concrete instantiations—entirely expressible in ordinary language. It is the most successful approximation of the semantic approach to truth prior to the 20th century.

2.3 Thomas Aquinas

The most famous statement about truth is doubtless:

(**TA**) Veritas est adequatio intellectus et rei (Truth is adequacy of thought and thing).

Everybody knows that this formula was given by Thomas Aquinas. However, (**TA**) covers only a part of Thomas’ definition, which has its full expression in (*De Veritate* I, 2).

(**TA1**) Veritas est adequatio intellectus et rei, secundum quod intellectus dicit esse quod intellectus dicit esse quod est vel non esse quod non est.

The difference between (**TA**) and (**TA1**) is not only in the length of both. The second part of (**TA1**) is simply Aristotelian explanation from *Metaphysics* 1011b (see (A3) in Chap. 1). It is rather obvious that Aquinas understood truth according to the Stagirite, and introduced the word *adequatio* as a convenient device to capture his basic intuition. He also employed such words as *assimilatio*, *conformitas* or *convenientia*, but always clearly noted that he followed Aristotle (see Schulz 1993 for an extensive treatment of Aquinas' philosophy of truth). In principle, Thomas repeats most Aristotelian convictions about truth, but modifies the Philosopher (as he addresses to the Stagirite) on one point. Namely, Aquinas stresses much more strongly the ontological aspect of truth. More precisely, he concludes one path (the second was taken by Duns Scotus, but I neglect this story) of the development of the theory of transcendentals (*transcendentalia* in Latin; see Knittemayer 1920, Schulemann 1929, Bärthlein 1972 and Aersten 1996 for more extensive accounts of the theory of transcendentals in ancient and medieval philosophy; the last book is mainly devoted to Thomas Aquinas). More specifically, Thomas argues that being, truth, good and one (some authors claim that also beauty belongs to this variety) are co-extensive in the following sense: if a and b are transcendentals, then for every x, x is a if and only if x is b. For instance, the thesis that *ens et verum convertuntur* illustrates the issue with respect to being and truth.

2.4 The History of the Adequatio Formula in the Middle Ages

Since the words used by Aquinas became decisive for the majority of later interpretations of truth as a correspondence relation, it might be interesting to look at how (**TA**) entered philosophy (see Gilson 1955, pp. 628–629, Boehner 1958, pp. 174–200). Thomas Aquinas himself notes that the adequatio formula appeared in *Liber de definitionibus* (the original Arabian title: *Sefer ha-Gewulin*) of Isaac Israeli (Honain ben Ishak), a Jewish historian compiler who died in 876. However, in the text (or rather in its Latin) translation we find only:

> Et sermo quidem dicentis: veritas est quod est, eniuntiativus est naturae veritatis et essentiae ejus quoniam illud sciendum quod est vera est.

This fragment simply does not contain the word *adequatio*, although certainly Isaac's statement can be understood as expressing an idea of correspondence. Avicenna's definition of truth in his *Metaphysica* I, 4 (Latin translation) runs as follows:

> veritas […] intelligitur dispositio in re exteriore cum est ei aequalitas.

We have here the word *aequalitas*, but not *adequatio*. Similarly, Averroes says (*Destructio destructionum* I, 3):

> Veritas namque, ut declaratum est in sua declaratione (definitione), est aequare rem ad intellectum scilitet quod reperiatur in anima sicut est extra animam.

The word *adequatio* appeared for the first time in the works of Philip the Chancelor, William of Auxerre (Guillemus Altisioderensis) and William of Auvergne, all of whom lived in the first half of the 13th century. Altisioderensis uses in his *Summa aurea* I, 1 the formulation (it is regarded as the oldest source with *adequatio*):

> Veritas [*est*] adequatio intellectus ad rem.

In *Summa de bono* I, 2, Phillip gives the following definition:

> Veritas [*est*] adequatio rei et intellectus, sive ut generaliter dicatur adequatio signi et significati,

which was later (but before Aquinas) almost literally repeated by Alexander of Hales in his *Summa Theologiae*:

> Veritas [*est*] adequatio rei et intellectus, sicut generaliter adequatio signi et significati.

It is possible that Philip the Chancellor was influenced by Abelard when he said that *signum* and *significandum* are involved in truth-definition; Bohner (ibidem, p. 180) notes similarity (also in the case of Alexander) with Anselm of Canterbury. William of Auvergne (*De universo* III, 1) says:

> […] et hoc [intentio veritas] ait Avicenna, est adequatio orationis et rerum; [veritas est] adequatio intellectus ad rem.

Thus, Wilhelm replaced *aequalitas* by *adequatio* in his explanation how Avicenna understood truth and then formulated his own definition using the words *adequatio intellectus ad rem.* Finally, let me note the definition given by Albertus Magnus (*De sententio* 46), the teacher of Aquinas:

> […] et quod [veritas] est adequatio rerum et intellectus, ut quando dicimus, ita intellectus est verus, referendo hoc ad sententiam intellectus conceptam de re sicut est.

This sample of quotations shows that the *adequatio* formula became very common prior to Aquinas. As far as the issues concerns the term *adequatio* and its practical functions, it seems that it could be a convenient technical word for truth-theory and, on the other, hand, easily memorized by students. However, it seems that in early Scholasticism there was competition between two views of the *adequatio* relation. One was semantic: *adequatio* expressed a semantic relation between *signum* and *significandum* (Anselm, Abelard, Philip the Chancellor, Alexander of Hales and, to some extent, Albertus; the last also refereed to semantic terms). The second view was more ontological and epistemological (Albertus Magnus in the quoted fragment; see Senner 1995 for an analysis of Albertus' view on the concept of truth).

I am not suggesting that medieval philosophers (perhaps Abelard was an exception) were aware that *adequatio* could be understood in two ways and it is possible that both understandings were present in the works of particular thinkers.

However, it seems that Thomas Aquinas took the second path and concentrated on epistemological and ontological issues related to the concept of truth. The historical record suggests that his view largely determined the later development of the *adequatio* formula and the correspondence theory of truth. Semantic considerations were quite intensive in the 14th century (in particular in William Ockham; see Perler 1992 for truth-theories in later Scholasticism), although this tradition did not influence later views, similarly as ideas of Abelard. That is very regrettable, because, for instance the Schoolmen's subtle reflections on the problem of how propositions are related to sentences turned out to be a philosophical novelty.

2.5 From Renaissance to Kant

I recall once again that my basic aim in the historical chapters of this book is to look for traces of the idea of correspondence as the fundamental concept of truth-theory. For this reason, I do not enter into comparative and interpretative problems pertaining to the truth-definitions sampled in this section. It is of the utmost interest that (**TA**) (see above) served as a very popular formula, often quite independently of a wider philosophical context of specific theories of truth. In particular, this general environment was either empiricist, like in Locke, or rationalistic, like in Descartes. In order to illustrating the actual impact of the idea of correspondence, I will list several typical statements about truth made by few philosophers from the 16th–18th centuries (translations are mine, except (12); the sequence of quotations is not exactly ordered by chronology):

(1) Goclenius, *Lexicon Philosophicum quo tantquam clave philosophiae fores aperiuntur*, p. 311:

> Veritas […] conformitas ut intellectae cum re existence extra intellectum;

(2) Gassendi, *Syntagma philosophiae Epicuri* I, 1:

> Veritas autem enuntiatonis seu iudicii nihil alliud est quam conformitas ore factae aut iudicii mente petracto cum ipsa enuntiata seu iudicata. (The truth of a sentence or proposition is nothing else than the conformity of a fact conceived in the mind and what is enunciated or asserted).

(3) Poinsot, *Ars Logica*, p. 274

> Veritas (iudicium) […] consistit in conformitate ad esse in conformitate ad esse vel non esse rei (The truth of a proposition consists in agreement with the existence or non-existence of the thing).

(4) Suarez, *Disputationes metaphysicae* 6, 2:

> Veritas transcendentalis significat entitatem rei, connotondo cognitionem conceptum intellectus, cui talis entitas conformatur vel in quo talis res representatur. (Transcendental truth means the being of thing which is connotes by knowledge when such entity conforms with represented thing).

(5) Descartes, "A letter to Mersenne" (1639):

[…] mot verité, en sa propre signification, denote la conformité de la pensée avec l'object. (The word truth in its proper meaning denotes the conformity of a thought with its object).

(6) Locke, *An Essay Concerning Human Understanding* IV, V, 9:

Truth is the making down in words the agreement or disagreement of ideas as it is […]. Signs […] contain *real truth* when […] are joined, as our ideas agree, and when our ideas are such as we know are capable of having an existence in nature but by knowing that such.

(7) Hume, *Treatise on Human Nature*, III, Part I, Section I:

Truth or falsehood consists in an agreement or disagreement either to the *real* relations of ideas, or to *real* existence and matter of fact.

(8) Wollaston, *The Religion of Nature Delineated*, Section I:

Those propositions are true which express things as they are; or truth is conformity of those words or signs by which things are expressed, to the things themselves.

(9) Spinoza, *Ethica*, axiom 6:

Idea vera debet convenire cum suo ideato (A true idea should conform to what is given in it).

(10) Leibniz, *Nouveaux Essais* IV, 5. 11:

Contentons nous de chercher la verité dans le correspondance des propositions qui sont dans l'esprit, avec le choses dont il s'agit. (Let us agree that we search truth in correspondence of propositions which are in mind with given things).

(11) Wolff, *Philosophiae rationalis sive logica* 505:

Veritas est consensus iudicii nostri cum objecto seu re representata (Truth consists in agreement of our proposition with the object which it represents).

(12) Kant, *Kritik der reinem Vernuft* A58; Eng. tr. (by N. Kemp Smith), *Immanuel Kant's Critique of Pure Reason*. London: Macmillan, 1929):

Was ist Wahrheit? Die Namenklärung der Wahrheit, dass sie nämlich die Übereinstimmung der Erkenntnis mit ihren Gegenstande sei, wird hier geschenkt und vorausgesetz. (What is true? The nominal definition of truth, that is the agreement of knowledge with its object, is assumed as granted; tr. by N. Kemp Smith, *Immanuel Kant's Critique of Pure Reason*, Macmillan: London 1929).

Suppose that someone without a deeper knowledge of the history of philosophy looks at (1)–(12). It is probable that such a person would say that all these authors offer us a similar idea, namely, that truth consists in a relation, called agreement, correspondence, similarity, etc. of something, for instance, sentence, proposition, idea, sign, mind, thought or knowledge—with something else, say fact, things, object, etc. And yet, the authors of the quoted definitions represented fairly different views concerning truth. In fact, only Goclenius, Gassendi, Poinsot, Suarez, Locke, Wollaston and Hume can be regarded as defenders of the *adequatio*-formula, although their general philosophical context was more or less different, even in the case of first four, who continued the Scholastic tradition. The very Cartesian account of truth is much better approximated by his famous statement that *verum est quod clarae ac distinctae percipio* which expresses the main tenet of the evidence theory (see Vinci 1998 for Descartes' theory of truth). Spinoza and Leibniz (for Spinoza's and Leibniz's aletheiologies, see Mark 1969 and Rauzy 2001) were more inclined to the coherence theory than to the idea of correspondence (see Walker 1989), both in different way. Spinoza derived coherentism from his holistic panteism, but Leibniz appealed to logical relations and sempiternal harmony of being. Although Wolff offered a typical correspondence formula, his definition was supplemented by remarks on how to investigate the content of propositions in order to achieve their coherence.

Kant is famous for his attack on the correspondence theory. His statement quoted as (12) only asserts that the agreement of knowledge and its object is simply obvious because object is constructed in the process of cognition and this fact precludes any disagreement between both. Kant continues (I give English translation only):

> The question asked [that is, what is truth? – J. W.] is as to what is the general and sure criterion of truth of any and every knowledge.

Thus, according to Kant, the real problem lies in truth criteria, not in the definition itself (see Scheffer 1993 for a penetrating analysis of this question). Yet Kant used extensively the word 'Übereinstimmung', so fundamental for aletheiology in the philosophical German. Other labels used in the period in question include 'agreement', '*conformitas*', '*conformité*', 'conformity', '*convenientia*', '*correspondance*' or '*oveerenkomende*' (the last word, occurring in Spinoza's early writings in Dutch refers to strict similarity; Jan Van Besten informed me that the etymology of *Übereinstimmung* probably is related to *oveerenkomende*. Returning to Kant, his views were close to coherentism, particularly in *Reflexionen Kants zur kritischen Philosophie*, ed. by B. Erdmann, Leibizg: Fues's Verlag 1882, p. 574, where he says that the entire truth consists in the agreement of all thoughts with laws of thinking, as well as in the mutual agreement of thoughts.

The above survey, although very selective and fragmentary, sufficiently confirms that the correspondence formula was used in the first period of modern philosophy as a convenient scheme for recoding various, often conflicting intuitions. Moreover, particular formulations employ fairly different words, very often without a proper care for precision. This observation concerns not only terms used for truth-bearers ('sentences', 'propositions', 'ideas', 'knowledge', all in various senses, especially, logical or psychological), but also such crucial words as 'representation', 'idea'or 'sign'. This may have been a result of the dominance of epistemology with a psychological flavour. Anyway, ontological questions pertaining to the concept of truth were decisively secondary in this period of the development of philosophical aletheiology.

References

Arsten, J. A. (1996). *Medieval Philosophy & the Transcendentals*. Leiden: Brill.

Bärthlein, K. (1972). *Der Transzendentalehre der alten Ontologie*, Teil I: *Der Transzendentallehre im Corpus Aristotelicum*. Berlin: de Gruyter.

Boehner, P. (1958). *Collected Articles on Ockham*. New York: The Franciscan Institute St. Bonaventure.

De Rijk, L. M. (Ed.) (1956). *Petrus Abaelardus Dialectica*. Assen: Van Gorcum.

De Rijk, L. M. (1956a). Introduction. In De Rijk (1956), I–XCVII.

Eggensperger, T., Engel, U. (Eds.) (1995). *Wahrheit. Versuche zwischen Hochscholastik Und Postmoderne*. Mainz: Matthias-Grünenwald Verlag.

Enders, M. (1999). *Wahrheit und Notwendigkeit. Die Theorie der Wahrheit bei Anselm of Canterbury im Gesamtzusammenhang seines Denkens und unter besonderer Berücksichtigung seiner antiken Quellen (Aristoteles, Cicero, Augustinus, Boethius)*. Leiden: Brill.

Gilson, E. (1955). *History of Christian Philosophy in Middle Ages*. New York: Random House.

Knittemayer, H. (1920). *Der Terminus 'transcendental' in seiner geschichtlicher Entwicklung bis zu Kant*. Mahrburg: Phil. Diss.

Mark, T. (1969). *Spinoza's Theory of Truth*. New York: Columbia University Press.

Perler, D. (1992). *Der Propositionale Wahrheitsbegriff im 14. Jahrhundert*. Berlin: de Gruyter.

Rauzy, J.-B. (2001). *La doctrine Leibnizienne de la vérité. Aspectes logiques et ontologiques*. Paris: Vrin.

Scheffer, T. (1993). *Kants Kriterium der Wahrheit. Anschauungsformen und Kategorien a priori in der "Kritik der reinen Vernuft"*. Berlin: de Gruyter.

Schulemann, G. (1929). *Die Lehre von der Transzendentalien in der scholastischen Philosophie*. Leipzig: Teubner.

Schulz, G. (1993). *Veritas est adequatio intellectus et rei. Untersuchungen zur Wahrheitslehre des Thomas von Aquin & zur Kritik Kants an einen überrlieferten Wahrheitsbegriff*. Leiden: Brill.

Senner, W. (1995). Zur Definition der Wahrheit bei Albertus Magnus. In Eggensperger, Engel (1995), 11–48.

Vinci, T. C. (1998). *Cartesian Truth*. New York: Oxford University Press.

Walker, R. C. (1989). The Coherence Theory of Truth. Realism, Anti-realism, Idealism. London: Routledge.

Chapter 3
Truth in the 19th and 20th Centuries

Abstract Views on truth of German and Austrian philosophers (among others, Lotze, Fries, Erdmann, Mauthner, Bolzano, Brentano, Frege) as well as British philosophers (Bradley, Joachim, Russell) are presented in this chapter. As Tarski as a mathematician and philosopher grew up in Poland, the special section is devoted to truth-definitions put forward by some Polish philosophers and logicians of the 20th century. The last section contains a sample of definitions circulating in philosophy in the last hundred years—most definitions are taken from analytic philosophy.

3.1 Introduction

This chapter addresses discussions pertaining to the correspondence theory of truth in the last 200 years. The material is even more selective than in the historical chapters above, due to the growing density of philosophical exchanges in the 19th and 20th centuries; moreover, is too early to include even a fragmentary historical report of the present period. That does not mean that everything of interest and importance to the philosophy of truth during this period is captured in the views reported in the following sections. For obvious reasons, I disregard German transcendendal idealism (Fichte, Schelling, Hegel), which does not appeal to analytic taste. Neo-Kantianism is also omitted even though, although philosophers belonging to this camp produced several influential ideas, above all, the concept of validity (*Geltung*)—which goes back Hermann Lotze. This idea was intended as a general category covering truth, good and beauty as its species, all existing in the transcendental realm. Validity and its kinds were conceived by the Neo-Kantians purely epistemologically. Such words as *gültig* or *Gültigkeit* used for 'true and 'truth' by German logicians, are instructive examples here. Yet they also functioned in German the broader philosophical vocabulary, not necessarily Neo-Kantian. One example of this influence will be provided in the next section. I am also very selective on Austrian tradition (see Simons 1992 for a wide panorama of this axis, including a part of the 20th century). Finally, let me note that I do not enter into various possible interpretations of views of particular philosophers, for example, Frege and Russell.

J. Woleński, *Semantics and Truth*, Logic, Epistemology, and the Unity of Science 45, https://doi.org/10.1007/978-3-030-24536-8_3

3.2 German Philosophers on the Correspondence Formula

The term *Übereinstimmung* became standard in German discussions on truth in the period in question. In fact, the idea of correspondence was used freely by most academic philosophers in Germany, and usually was regarded as entirely obvious. I will give five examples from German philosophical literature:

(1) Fries 1837, p. 308:

> Nennen wir die Übereinstimmung einer Erkenntnis mit ihren Gegenstand, ihre transzendentale oder Vernuft-Wahrheit (Transcedental or intellectual truth of knowledge is called its conformity of knowledge with its object).

(2) Erdmann 1892, p. 367

> Die Definition der Wahrheit im eigentlichen Sinne als Allgemeingültigkeit geht […] auf die Aristotelische zurück.
>
> (Definitions of truth in the proper sense as universal validity goes back to Arisstotle).

(3) Mauthner 1902, p. 360

> Die Wahrheit unserer Erkenntnis ist die Übereinstimmung unserer Urteile mit der Wirklichkeitswelt; da unsere Urteile rückschreitend bis auf Sinnesausdrücke so ist die Wahrheit unserer Erkenntnis schliesslich auch die Übereinstimmung unserer Vorstellungen und Sinnesausdrücke mit der 'Wirklichkeit'. (Truth of our knowledge consists in agreement of our judgments with the real world; since our judgments are reducible to sense-expressions, truth of our knowledge is finally reducible to conformity of our presentations und sense-expressions with 'reality'.)

(4) Meier 1926, p. 223

> Ungesucht bietet sich die alte *aristotelische* Antwort dar, die bis in die Gegenwart herein ihr Ansehen behauptet hat: das Urteil [ist] wahr […]: es will mit der Wirklichkeit übereinstimmen. Die Unhaltbarkeit dieser Definition fählt indessen in die Augen, sobald man ihr nun ihre genauer Fassung, gibt. Nicht von einer Übereinstimmung des Urteils, sondern nur von einer Übereinstimmung des Urteils *gegenstandes* mit der Wirklichkeit kann die Rede sein. In der Tat ist dies der genuine Sinn der aristotelischen Wahrheitstheorie. (A simple Aristotelian answer is popular until the contemporary period: a judgment is true if it agrees with reality. The untenability of this definition became obvious if we consider it more closely. One cannot speak about the agreement of a judgment with reality, but only about agreement of objects of judging with reality. In fact, it is the genuine meaning of Aristotelian truth-theory.)

(5) Eisler 1930, p. 450/451

> *Materiale* [Wahrheit] ist, ganz allgemein, "Übereinstimmung" (Konformität) des Denkens mit dem Sein. Es gibt aber zwei Arten der materialen [Wahrheiten]: (a) *Empirisch-immanente* […]. Hier bedeutet "Übereinstimmung" von Denken und Sein […] *nicht* die

> *Abbildung* u. dgl. des Seienden im und durch das Denken, sondern *Übereinstimmung* des *Einzelurteils mit der methodisch gesetzten Realität*, die in einen System von Warhnemungs-und-Urteilsnotwendigkeiten sich darstellt […]. (b) *Metaphysische* [Wahrheit] ist die Übereinstimmung des Denkens mit der absoluten Wirklichkeit […]. Auch hier kann von einen "Abbilden" keine Rede sein, sondern die "Übereinstimmung" bedeutet hier ein mehr oder weniger treffendes "Nachkonstruieren" der transzendenten Wirklichkeits-Verhältnisse *in immanenten*, *begrifflichen Symbolen*. (Material (truth) is, quite generally, "agreement" (conformity) of thought with being. There are two kinds of material truths: (a) *empirical-immanent* […] "Agreement" of thought and being means here *no* picture of what is being by thought, but *agreement* of *a particular judgment with reality which is established methodically* and presents itself by a system of empirical and judgemental necessities […]. *Metaphysical* [truth] is agreement of thought with absolute reality […]. Also here one cannot say about any "picturing", but "agreement" means in this case more or less deep "construction" of transcendent relations within reality by *immanent*, *conceptual symbols*.)

One could find hundreds of similar answers to the question "What is truth?" They occurred in textbooks (like (1), (3)), scholarly monographs ((like (2), (4)) or dictionaries (like (5)). Each of these (and similar) quotations could be the topic of an extensive philosophical essay. Consider (2), for example, Benno Erdmann tried to defend Aristotle by using the mentioned idea of validity. On the other hand, he radically rejected the Scholastic reading of the Stagirite via such labels as *adequatio*, *conformitas* or *assimilatio*, because truth is objective and cannot be explained by referring to concrete mental acts. I consider Erdmann's interpretation of Aristotle completely mistaken. In fact, it would be very difficult to find a way to justify an interpretation of the formula 'to say that that which is, is and that which is not, is not, is true' via *Allgemeingültigkeit*. Thus, Erdmann over-interpreted Aristotle by using German philosophical terminology of his time. Also, his criticism of the *Übereinstimmung* (*adequatio*, *conformitas*, *assimilatio*) concept seems oversimplified by the philosophical background he invokes. Turning to (1), (3), (4) and (5), all of these formulations employ a considerably rich and complicated vocabulary, and provoke objections similar as in the case of (2). I offer only two remarks with respect to these definitions (or rather explanations). The first is historical. In (4) we read that this explanation displays the original meaning of Aristotle's theory of truth. However, Aristotle, who was a strong realist, could not say that truth consists in agreement between reality and objects of judgments (propositions), even if he were use the term 'agreement', he did not propose that truth consists in a relation. Thus, (4) proposes a very arbitrary reading of Aristotle. Second, I am inclined to think that such complicated explanations of the idea of correspondence as occurs in (1)–(5) contributed essentially to the aversion of some logicians and philosophers toward the idea of truth as a correspondence (see below). Although I am personally very far removed from attributing the formula *veritas est adequatio intellectus et rei* to the Stagirite, I think that the issue is much deeper than (1)–(5) might suggest.

3.3 The Concept of Correspondence in British Philosophy About 1900: Bradley, Joachim, Russell, Moore

Two debates on the concept of truth occurred in British philosophy at the end of 19th century and the beginning of the 20th centuries: one involved the pragmatic account, and the other was devoted to coherence and correspondence as fundamental categories in explaining the nature of truth. Francis Bradley, Harold Joachim and Bertrand Russell collectively argued against Ferdinand Schiller and William James as pragmatists, but, on the other hand, Russell and G. E. Moore defended the correspondence theory against Bradley and Joachim as coherentists (see also Bosanquet 1911, v. 2; I neglect here differences between representatives of this approach). My objective in this section is to present the second debate to the extent that the concept of correspondence is involved in it.

(**DG1**) It is perhaps worth noting that James (see James 1907, p. 57) regarded the formula that truth consists in the agreement of a judgement with reality as obvious, but immediately added that the nature of this relation is controversial. On the other hand, assimilation, validation, corroboration and verification are not controversial as marks of truth (p. 201). Of course, they function as criteria of usefulness. James' remarks show, on the one hand, that the correspondence theory of truth can be explicated in many ways, but, on the other side, that the issue deserves to be carefully analysed in concrete philosophical idioms (see Sprigge 1997 for a comparison of James and Bradley and a detailed expositions of truth-theories of both philosophers, and Nesher 2002 for a general picture of aletheiology from the pragmatist point of view—this book has several historical analyses, for instance, Spinoza's views on truth.►

The concept of correspondence was introduced into British philosophy not by a professional philosopher, but Samuel Coleridge, a poet who applied (in 1809) the term 'correspondence' in the context of a theory of truth (quoted after *The Oxford English Dictionary*):

> By verbal truth we mean the correspondence of a given fact to given words.

However, the word 'correspondence' had to wait almost a hundred years to become popular among professional philosophers. Then, this term was occasionally reintroduced by Bradley, probably independently of its usage by Coleridge (Bradley 1883, p. 551; page-reference to Bradley 1922):

> The validity of inference has two main senses […] We might ask if in argument we possess a strict counterpart of the nature of things, if our mental objects truly represent any actual process […] And this would be the first question. The second would ignore this correspondence with reality.

When Bradley was speaking about correspondence with reality as applied to truth, 'truth as copying' was his favourite label (see the title of Bradley 1907).

(**DG2**) The word *correspondentia* did not occur in classical Latin. It appeared in the Middle Ages. It had two meanings: (a) conformity, similarity; (b) gratitude, reciprocity. Omitting (b) (it does, however, have a moral aspect), (a) has its genesis in the verb *correspondeo*—agree, answering, etc. The word *respondeo* (give an answer) could be regarded as another link. Hence, *co-respondeo* involves at least two agents exchanging views, although not necessarily being in agreement. A further suggestion takes into account *spondeo*, that is, promising something with intention to fulfil what is promised. Anyway, *correspondentia* in the meaning (a) seems an abstraction from a more primitive ordinary usage. Entering this word into the philosophical parlance required time. As far as I know (I am not a linguist) the Schoolmen did not use *correspondentia* in their explanations concerning truth. Even if I am mistaken, this word appeared very rarely in philosophical writings of the Middle Ages. It seems that Leibniz was one of the first to apply the French counterpart of *correspondentia*, that is, the word *correspondance* (see Chap. 2, Sect. 2.5(10)).►

It was Harold Joachim who was responsible for making the term 'correspondence' standard in the contemporary philosophical English. He opened his book on truth by the following sentences (Joachim 1906, p. 7):

> In most of the everyday judgements of common sense, and in many philosophical theories, a certain conception of truth is implied or expressed, which I shall call the 'correspondence-notion' of truth. Thus, e.g. to 'speak the truth' is to speak 'in accordance with' or 'in conformity to' the facts.

However, Joachim's mentioned book on truth was influential for more important reasons than terminology. Bradley and Russell (see Candlish 2007 for the Russell/Bradley debate; see also Woleński 1994), other *dramatis personae* in this controversy, did not contribute very much to the correspondence theory of truth before Joachim's book appeared. A criticism of the 'truth as copying' clearly resounds in Bradley 1883, pp. 579–580 and the notes added in Bradley 1922, p. 592.

Early Bradley was not particularly interested in a criticism of the 'copying theory'. His main objective in the theory of truth consisted in developing of coherence theory within the system of monistic metaphysics; he did it in Bradley 1893. Russell tried to reduce truth to propositions (I choose this category of bearers as basic disregarding other Russell's wordings) in the logical sense (see Russell 1903, p. 48, p. 504), and pointed out that there is a very serious difficulty in explaining of propositions refer to facts. The actual problem is this. Assume that we ask what is asserted by the proposition 'Caesar died'. A plausible answer is: the death of Caesar. Now, on the one hand, this seems to imply that just the death of Caesar is qualified as being true or false, but on the other hand, logical values, that is, truth and falsehood, not being logical subjects. And Russell continues (Russell 1903, p. 48).

> The answer here seems to be that the death of Caesar has and external relation to truth of falsehood (as the case may be), whereas "Caesar died" its own truth or falsehood as an element. But if this is the correct analysis, it is difficult to see how "Caesar died" differs from "the truth of Caesar's death" in the case where it is true, or "the falsehood of Caesar's death" in the other case. Yet is quite plain that the latter, at any rate is never equivalent to "Caesar died." There appears to be an ultimate notion of assertion, given by the verb, which is lost as soon as we substitute a verbal noon, and is lost when the proposition in question is

> made by the subject of some other proposition. This does not depend upon grammatical form; for if I say "Caesar died is a proposition," I do not assert that Caesar did die, and an element which is present in "Caesar died" disappeared. Thus the contradiction which was to have been avoided, of an entity which cannot be made a logical subject, appears to be inevitable. This difficulty, which seems to be inherent in the very nature of truth and falsehood, is one with which I do not know how to deal satisfactorily.

However, Russell may have not considered this difficulty as particularly important, because one year later he wrote (Russell 1904, p. 523):

> It may be said – and this is, I believe, the correct view – that there is no problem at all in truth and falsehood; that some propositions are true and some false, just as some roses are red and some white; that belief is a certain attitude toward propositions, which is called knowledge when they are true, error when they are false.

The situation changed radically after Joachim's book. Russell reviewed it twice in 1906 (see Russell 1906, Russell 1906a). Then, he published an extensive criticism of Joachim together with his own account of what is truth (see Russell 1907). This same year brought Bradley 1907, Moore's detailed review of Joachim 1906 (see Moore 1907), and Joachim's reply to Moore (see Joachim 1907). In 1910, Russell 1910 appeared with two essays on truth. Russell 1912 was published two years later with a separate chapter on truth. Several essays of Bradley, are collected in Bradley 1914; this book includes already published, for instance Bradley 1907, as well as new papers. Moreover, Moore addressed himself to the discussed problem in his lectures of 1909–1910, published more than forty years later (see Moore 1953). Russell's extensive treatise on the theory of knowledge, completed in 1913, but only published as Russell 1984 also contains a chapter on truth. This survey clearly shows that Joachim's book initiated a very spirited discussion on the concept of truth. Of course, the story goes beyond 1914, but its subsequent course exceeds the scope of this work. I also omit here Moore's defence of the correspondence theory, because he simply assumes that this theory is fairly obvious from the common-sense point of view.

Joachim devotes two chapters of his book of 1906 to the concept of correspondence and problems connected with it. Chapter 1 ("Truth as Correspondence") gives a reconstruction of the correspondence-concept as well as it offers a criticism of the theory based on this notion; Joachim, along with many other authors, attributes this account of truth to Aristotle. Chapter 2 ("Truth as a Quality of Independent Entities") criticizes the approach expressed in Russell 1903 which attributes truth or falsehood to assertions in the logical sense, regarded as entities that are independent of particular human minds. One should see the actual historical significance of Chap. 2 in provoking Russell to elaborate his own correspondence theory as contained in Russell 1907, Russell 1910 and Russell 1912. Since Joachim says in the preface to his book that Russell read Chap. 2 before the book has appeared, it is probable that he regarded Joachim's *The Nature of Truth* as a serious challenge.

(DG3) Here are two of Russell's truth definitions:

(a) Russell 1910, p. 156:

> Every judgement is a relation of mind to several objects, one of which is the relation; the judgement is true if the relation which is one of the objects relates to the other objects, otherwise is false.

(b) Russell 1984, p. 144:

> The belief is true when the objects are related as the belief asserts that they are.
>
> Thus the belief is *true* when there is a certain complex which must be a definable function of the belief, and which we shall call the *corresponding* complex, or the *corresponding* fact.

Both—explicitly as in (b) or implicitly as in (a)—explain truth via using the concept of the corresponding fact. Russell's role in the history of alethoiology is much greater than his polemical exchanges with Bradley and Joachim or his attempts to define truth, because his theory of logical types plays a fundamental role in the foundations of mathematics and semantic considerations in the years 1920–1939, but I skip this topic. See also Russell 1940, Chap. XVI.►

Returning to Joachim, the correspondence always involves two constituents that are connected by a structural similarity or some other one-to-one relation. Thus, if one claims that the nature of truth consists in correspondence, one must define both members of the correspondence-relation. Of course, this relation holds between something mental and something factual. Moreover, the correspondence in question must be comprehended by a mind. However, as Joachim argues, it is impossible to separate the real from the mental or the mental from the real. This is a straightforward consequence of the observation that the correspondence–relation holds 'for a mind'. Thus, it is false to say that the correspondence-relation holds between a mental factor and a mind-independent reality: this relation cannot be purely external. At the same time, Joachim notes (Joachim 1906, p. 20):

> Truth […] has its own stubborn nature to which our thinking must conform on pain of […] error. We do not make or alter truth by our thinking […]. Truth is discovered, and not invented; and its nature is not affected by the time and process of discovery […]. It is to this independent of entity that the judgement of this or that person must conform if he is to attain truth. Correspondence of his thinking with this 'reality' is truth for him; but such a correspondence requires an independent truth […] as one of its factors and is not the essence of truth.

Joachim inherited the views expressed in this quotation from Bradley. Thus, we see that British Neo-Hegelians admitted that a correspondence is essentially involved in the concept of truth. However, this is quite different concept of correspondence than that used by the typical correspondence-theoreticians, because the relation in question is internal and derivative with respect to a more basic feature of truth—that is, coherence.

Bradley himself decided to analyze the correspondence theory of truth (the theory that truth is copying). His main argument against this theory is following (Bradley 1907, p. 108):

> […] the whole theory goes to wreck in principle and at once on a fatal objection. Truth has to copy facts, but on the other side the facts to be copied show already in their nature the work of truth-making […] much of given fact is inferential […] if there really is any datum, of outward or inward, which if you remove the work of the mind, would in its nature remain the same, yet there seems no way of our getting certainly to know of this. And, if truth is to copy fact, then truth at least seems to be in fact unattainable.

Roughly speaking, thought and reality belong to different ontological realms and cannot be compared. Once again, because truth consists in the holding of internal relations, it cannot be explicated by referring to the external relation of comparing items existing at different levels of reality. Although this argument is relatively independent of Bradley's holistic metaphysics, the rest of Bradley's objections against the correspondence theory is essentially based on his metaphysical views, and must be skipped here. Let me just note again that a thesis that external relations (connexions) are impossible is basic in this context.

This last point is central for Russell (Russell 1910, p. 139):

> The doctrines we have been considering [those of Joachim – J. W.] may all be deduced from one central doctrine, which may be expressed thus: 'Every relation is grounded in the nature of the related terms.' Let me call this the axiom of internal relations.

I will not assess whether Russell's formulation of the axiom of internal relation is correct or sufficiently clear. What is important here is that Russell's own version of the correspondence theory consists in regarding the relation of correspondence as an external relation. Russell looks at truth as a property of beliefs or judgements (I do not enter into details of Russell's notions of judgment and belief); on the other hand, truth can also be attributed to sentences which express beliefs, but sentences are true in a secondary sense, that is derivative from beliefs being true. For Russell, truth does not consist in a relation to a single object "which is what we judge or believe" (Russell 1910, p. 150). Thus, if one truly judges that Caesar died, this truth does not consist in a relation of this judgement to the object denoted by the expression 'the death of Caesar'. On Russell's view, a judgement is a multiple relation between a subject and several objects and involves essentially what is judged. For example, if a person *P* judges that Caesar died, the relation of judging holds between *P*, Caesar and the dying of Caesar. Now, if a relation holds between Caesar and his dying, and is involved into the judgement of *P* that Caesar died, is the same as the actual relation which holds between Caesar and his dying, the related judgement is true, otherwise, it is false. It is important to note that Russell's account of the correspondence relation has nothing to do with copying or displaying facts by mental entities. In Russell 1912, Chap. 12, he states his famous requirements for any correct theory of truth: (a) the theory of truth must also explain the nature of falsehood; (b) truth is a property of beliefs; (c) truth consists in an external relation of beliefs to something existing outside them. Clearly, Russell excludes the coherence theory in advance.

Can we derive something interesting for our present discussions on truth from the Joachim–Bradley/Russell–Moore controversy over the notion of correspondence? I think that the answer is definitely: yes. The problem of external/internal relations, as discussed by the mentioned philosophers, can be, at least, partly restated as follows: is a purely extensional account of relations sufficient for the correspondence theory of truth? The suggested answer I submit is: not. The correspondence theory of truth seems to require relations-in-intension or at least intensions as meanings of expressions (see Chap. 7 on **STT** in this respect). Does that mean that we should return to Bradley and Joachim in the sense that introducing internal relations is indispensable for analysing the concept of truth? I do not think so. Since bearers of truth have meaning (or are meaningful), a combination of meaning and relations-in-extension captures the intensional aspect of relations, that is, associated with meanings of phrases expressing relations. This route was just taken by Tarski in **STT**. I will argue (see Chap. 9, Sect. 9.2) that this moment has a relevance for a satisfactory analysis of the correspondence concept from the semantic point of view.

3.4 Bolzano

Bolzano was a predecessor for many important ideas in logic and semantics. His *opus magnum* (Bolzano 1837, Sects. 24–34; (see also Gotthardt 1918)) contains general remarks on the concept of truth (I omit the more specific explanations in Sects. 198–221—they concern the role of truth in inferences). At first, Bolzano distinguished several meanings of the words *wahr* (true) and *Wahrheit* (truth):

(a) As property of propositions in themselves (*Sätze an sich*);
(b) As true proposition in itself;
(c) As assertion of a true propositions;
(d) As the totality of truths;
(e) ‘True’ as synonymous with ‘actual’ or ‘real’;

For Bolzano the sense (a) appears to be central. Propositions in themselves are abstract entities, fully independent of individual mental acts. Consequently, propositions in themselves should be considered as proper bearers of truth. Truths in this understanding do not exist in time and space—they are non-temporal and non-spatial. This also means that truths are not created by human beings. God is the only subject who knows all truths. These statements about truth entail that it is absolute, even if the theological environment of this thesis is dropped. Since propositions in themselves are Platonic-type objects, their properties must be continuously stable.

Bolzano’s account of truth, the most relevant in the context of **STT**, might be summarized by the following definition:

(**Bo**) The proposition A is true if and only if A attributes to a thing a a property P which is actually possessed by a, Assume that A has the form Pa. Consequently, Pa is true if and only if a actually possesses a property P. In fact, the definition (**Bo**) is very close to **STT** (the case of atomic sentences). Moreover, Bolzano's semantic notion of variation anticipates the concept of satisfaction (see Casari 2016) employed by Tarski in the construction of his truth-definition. On the other hand, the clause 'actually possesses' points out that Bolzano tried to combine semantic and ontological features of truth; the latter had importance for his view that truth is absolute. Bolzano did not use the term 'correspondence', but he did point out his affinity to Aristotle. As a declared anti-Kantian he criticized Kant's fundamental view that truth is construed by a mind equipped with a priori categories.

3.5 Brentano

Franz Brentano concentrated his considerations on truth around the formula *veritas est adequatio intellectus et rei* (see Brentano 1930; this book collects Brentano's papers on truth written, but mostly unpublished, during his lifetime; Srzednicki 1965 contains an extensive discussion of Brentano's aletheiology). For Brentano, this account appealing to this formula needs to be purified from various misinterpretations. In particular, he protested against the idea that truth consists in agreement with reality. This criticism was dictated by Brentano's account of bearers of truth as acts of judging. A judgement as an act asserts or rejects the existence of its subject. Hence, linguistically speaking, judging is essentially associated with affirming or denying the sentence 'a exists'. Consequently, the word 'correspondence' (or its aletheiological synonyms) can mean that truth consists in asserting an object if it exists or rejecting it if it does not exist. Conversely, judging is false, when it asserts what does not exist or rejects what exists. This account is essentially based on Brentano's view that every judgement is existential and that every mental act is intentional (directed to something, existing or not). Furthermore, as Brentano argues, correspondence cannot be defined as similarity relation or identity relation. Why did Brentano choose judgments as bearers of truth? His early answer pointed out that this approach simplifies truth-theory. In his later phase, Brentano as a reist, that is a philosopher maintaining that only singular objects, physical or mental, exist in the ontological inventory, argued that analysing the concept truth via acts of judging allows to dispense with *entia rationis* (abstract objects) as realities corresponding to intentional acts.

The later Brentano was not satisfied with the just outlined answer to the question 'What is truth?' Although he retains the intentionality thesis, he replaced his earlier definition of truth, by the following explanation:

(**Br**) If an object a exists, someone who affirms a, judges correctly, and if a does not exist, someone who rejects a, judges correctly.

According to Brentano, (**Br**) offers a proper interpretation of the *adequatio*-formula and, in particular, defines truth as absolute. Yet this view leads to the question of what makes judging correct. Brentano answers that correct judging must be evident. Now, evidence cannot be a property of judgments about physical things. By contrast, only inner experience (*innere Wahrnemung*) provides evident judging. I will not enter into the details of Brentano's notion of evidence, except two remarks—one systematic and one comparative. First, it is not quite clear, whether Brentano had in mind a property of a true judgment or a criterion of truth. Second, Brentanian evidence characterizes (or not) a kind of empirical knowledge and, thereby, evidence different than occurring in Platonic or Cartesian approaches to cognition. Returning to (**Br**), one can correctly say that the idiom 'judges truly' appears as a fairly proper expression of Brentano's ideas on the concept of truth. Thus, the adverb 'truly', and not the noun 'truth'—or even the adjective 'true' as referring to a property, appears as the main linguistic device of the Brentano-style aletheiology. This account is sometimes labelled as the adverbial theory of truth. Brentano is also occasionally interpreted as a deflationist (see Brandl 2017).

Brentano also formulated several objections against the notion of correspondence:

(a) The correspondence definition of truth does not rebut the arguments raised by sceptics and relativists;
(b) The correspondence theory of truth cannot explain what corresponds to negative judgements, in particular, to negative existential ones;
(c) The correspondence theory of truth does not explain why theorems of logic and mathematics are true, because such statements do not correspond to specific objects, but are universally valid;
(d) The correspondence theory of truth cannot rebut the circularity, infinite regress or petitio principii objections;
(e) The notion of correspondence is too vague in order to constitute a satisfactory foundation of a truth-theory.

Point (d) calls for a special explanation. Assume that the sentence A is true in the virtue of holding a suitable correspondence relation C_0. So we have A and 'A is true'. Let the symbol A_1 abbreviate 'A is true'. If A is true, A_1 is true as well in virtue of C_1. Moreover, A cannot be true without A_1 being true. This reasoning is repeatable with respect to A_2 asserting that A_1 is true, and continued ad infinitum. In other words, we have an infinite sequence $C_0, C_1, C_2, \ldots$ of correspondence relations. In order to justify that C_{n-1} holds, we need to confirm that C_n obtains. If we interrupt our explanation at C_n, this results in petitio principii or circularity, and if will continue, we fall into a regressus ad infinitum. Every satisfactory truth-definition must rebut objections (a)–(e).

3.6 Frege

According to Frege, truth is not definable. He says (Frege 1918, p. 353):

> […] explanation of truth as correspondence breaks down. And any other attempt to define truth also breaks down. For in a definition certain characteristics would have to be specified. And in application to any particular case the question would always arise whether it were *true* that the characteristics were present. So we should be going round in a circle. So it seems likely that the content of the word ‘true’ is *sui generis* and undefinable.

On the other hand, Frege considered the meaning of the adjective ‘true’ as perfectly clear and not requiring further explanations. The quoted fragment directly blames the correspondence definition of truth for its circularity, but Frege indirectly applied this objection to any other theory that attempts to define truth by means of an apparently simple category as, for instance, coherence.

Frege’s undefinability thesis does not mean that he did not formulate substantial assertions on truth (see Greimann 2007, Pardey 2012 for detailed presentations of Frege’s truth-theory). In particular, he was interested in the relation between truth and logic (I omit other truth-themes on which Frege commented). According to him (Frege 1918, p. 350):

> Just as ‘beautiful’ points the ways for aesthetics and ‘good’ for ethics, so do words like ‘true’ for logic. […] it falls to logic to discern laws of truth.

The following ideas are basic (see Frege 1979, pp. 1–8, 126–151, 174–175)

(a) Logic assumes the distinction of truth and falsehood;
(b) The True and the False as logical values are references of sentences;
(c) Thoughts (propositions) are senses of sentences (Frege’s famous distinction of *Sinn* and *Bedeutung* for sentences);
(d) Logic is the science of truth—it develops the meaning of the adjective ‘true’;
(e) Although logic is the science of truth, it does not collect empirical truths; this task belongs to the special sciences;
(f) Logic discovers the principles of correct inference, that is, of inferences suitable for justification of sentences on the basis of other sentences;
(g) An inference is correct independently of whether its premises are true or not;
(h) The truth of a conclusion of a correct inference depends on the truth of its premises;
(i) Only true sentences (Frege says: thoughts) can serve as premises;
(j) One cannot understand false sentences, provided they are asserted.

The thesis that logic is the science of truth should be understood to mean that logic concerns formal principles of truth (the distinction between formal truth and material truth was very common in German philosophy of the 19th century). Putting this, in a more contemporary terminology, logic investigates principles of logical entailment (logical consequence) as codified by logical calculus. According to Frege, logical entailment is a relation, which, if it holds, always guarantees the truth of the conclusion, if the premises are true; technically speaking, logical entailment preserves truth, that is, transmits it from premises to conclusion, provided logical entailment holds.

Since the True and the False are common references of all true or false sentences, logical laws operate on logical values and are independent of senses. An apparent conflict between (h) saying that an inference is correct independently of the logical value of its premises and (i) requiring its premises to be true, has an easy solution by introducing the distinction between formally correct inferences (their conclusions follow logically from their premises) and materially correct inferences (these are not only formally correct but also have true premises). Thesis (h) raises serious interpretative problems. Frege defined logical assertion in such a way that only truths could be asserted. Accordingly, asserting falsehoods is logically impossible, because that would amount to a contradiction. However, it can be said against Frege that one should distinguish logical assertion and psychological assertion (other terminology will be recommended below), analogically to the distinction of propositions in the logical sense and propositions in the psychological sense. Now, the impossibility of asserting a falsehood logically does not entail that asserting falsities psychologically cannot happen. The distinction of the two kinds of assertion is not at odds with Frege's view, because he merely claimed that logical and psychological matters should be sharply separated.

Frege's view about assertion, truth, and falsehood seems to stem from his denial of the definability of truth, because he considered logic as developing the content of 'true' by means of logical principles. This can be understood in two ways. First, logic develops the content of logical truth. In this interpretation, logical assertion concerns only principles of logic, that is, logical theorems. Since they asserted unconditionally, logical falsehoods cannot be asserted, because it is easily demonstrated that an assertion of a logical falsehood must be inconsistent. Let $\mathbf{As}(A)$ means 'A is asserted'. Assume two simple principles (in the contemporary setting on which **As** functions as a propositional operator):

(1) (a) $\mathbf{As}(A \wedge B) \Leftrightarrow \mathbf{As}(A) \wedge \mathbf{As}(B)$;
    (b) $\mathbf{As}(\neg A) \Rightarrow \neg\mathbf{As}(A)$.

Assume now $\mathbf{As}(A \wedge \neg A)$. Applying (1a) we obtain $\mathbf{As}(A) \wedge \mathbf{As}(\neg A)$, and (1b) immediately gives a contradiction in the form of $\mathbf{As}(A) \wedge \neg\mathbf{As}A)$.

It is uncertain whether Frege restricted his views about logical assertion to tautologies. Note above all that the phrases 'the logical content of truth' and 'the content of logical truth' do not need to be equivalent. It seems that Frege, in speaking about developing truth by means of logic was speaking about truth *simpliciter*, but not about logical truth. This brings us to a certain interpretation of Frege's view on assertion. This interpretation assumes that asserting falsehoods is inconsistent. Yet the outlined reasoning based on (1) does not apply to empirical falsehoods, unless one presupposes that false sentences are not asserted at all. But this last statement is empirically false, because asserting falsehoods is notoriously common. I think that the soundness of Frege's view on logic requires that assertion be replaced by assertibility. The claim that falsehoods cannot be rationally assertible seems reasonable. Accordingly, assertibility—not assertion—is a logical notion, and the phrase **As(***A***)** should be read '*A* is assertible' (see Picardi 1981 for an analysis of Fregean views on truth and assertibility, and Chap. 4, Sect. 4.8). Finally, Fregean assertibility is basically different than proposed by constructivists (for instance, intuitionist) as the base for anti-realistic conception of truth (see Chap. 9, Sect. 9.9).

Frege criticized the correspondence theory of truth using his idea that all true sentences refer to the True as their object. Assume that truth a proposition consists in its correspondence with its object. If so, all truths would correspond with the same objects. The contemporary version of this argument points out that true propositions correspond with the Great Fact, provided that the correspondence relation holds between truth-bearers and facts (the name 'the Great Fact' was proposed in Davidson 1969, but this argument was also formulated in Church 1943 and Gödel 1944; see also Neale 2001 for a general analysis of this issue; the slingshot argument is another label for this puzzle). I will briefly return to this issue in Chap. 9, Sect. 9.2.

3.7 Polish Logicians and Philosophers on Truth

3.7.1 Twardowski

Here is Tarski's testimonial (Tarski 1992, p. 20)

> Almost all researchers, who pursue the philosophy of exact sciences in Poland, are indirectly or directly the disciples of Twardowski, although his own work could hardly be counted within this domain.

Twardowski studied with Brentano and shared many of his philosophical views, including those concerning the concept of truth. Just like his master, Twardowski had very serious reservations about the *adequatio*-formula (see Twardowski 1975) and pointed out that this approach is based on metaphysical assumptions that are definitely too strong and should be very carefully investigates. Twardowski's treatment of the problem of truth-absoluteness became particularly important (see Twardowski 1900) for aletheiology developed in Poland. Something (a truth-bearer) is absolutely true if it is true everywhere, at all times and under all conditions. By contrast, something is relatively true if it is true at some places, sometimes or under some conditions. Twardowski intended to show that there are no relative truths. He considers several examples of apparent relative truths, for instance, statements that such and such flower has a pleasant scent, that cold baths are healthy, that it is raining, or that empirical hypotheses are only temporary. Twardowski argued that these formulations are inexact and require a further analysis.

Twardowski claimed that one should sharply distinguish sentences (*powiedzenia* in Polish) and propositions (*sądy* in Polish).

(**DG4**) There is a problem how to translate the word *sąd* into English in order to faithful for Twardowski's intentions. For him, judging is a mental act with judgment as its product. Thus, the term 'judgment' seems proper in this context. On the other hand, the term 'proposition' better fits contemporary discussions, because it is less psychological, and I will use it that way. Let me add that the act/product distinction is absolutely fundamental for Twardowski (see Twardowski 1912) and the entire Lvov–Warsaw School—particularly for the philosophy of language.►

Propositions are proper bearers of truth, but sentences are true or false in a derivative sense. Now, sentences are frequently incomplete and require supplementing by additional information. Consider the sentence (a) 'It's raining'. It so happens that the context decides whether (a) is true or false. However, this sentence does not express a proposition. The situation changes if we convert (a) to (b) 'At 12 noon, Central European time on March 1, 1900 according to the Gregorian Calendar, it is raining in Lvov on the Akademicka Street' (Twardowski's original example is slightly different). Given that (b) is true, it can be false if asserted, for example, about the Main Square in Cracow at the same time. A similar analysis enables us to demonstrate that other examples provided by relativists (Twardowski regarded pragmatism as a typical relativism) can be converted into absolute truths. As far as the issue concerns empirical hypotheses, Twardowski claimed that we should distinguish truth and our knowledge that a sentence is true or false. Briefly speaking, sentences can contain indexicals or other relativizations, for instance, to taste or the stated of knowledge, but propositions are always true or false. Moreover, Twardowski pointed out that relativism is at odds with basic logical rules, like the principles of contradiction and the excluded middle.

3.7.2 The Kotarbiński–Leśniewski Debate in 1913

Tadeusz Kotarbiński (see Kotarbiński 1913) considered the problem of the existence of the future. His starting point was the following truth-definition (Brentanian in its essence):

(1) A proposition A affirming an object a is true if and only if a exists.

He also accepted

(2) For every A, if A is true at time t, it is also true at every t_1 later that t.

This last statement expresses that truth is eternal. On the other hand, Kotarbiński rejected

(3) For every A, if A is true at t, it is also true at every t_1 earlier than t,

that is, the sempiternality of truth. Some truths are eternal and sempiternal, but other cannot be qualified in such a way. To show that let us consider an object a created by a human being at t. Since a does not exist until it gets created, no proposition A is true about it. But A is not false either before t, because its negation would be eternally true and, hence, it would be impossible to create *a* at any time, contrary to the assumption that the object in question was created at *t*. Consequently, *A* is not a sempiternal truth and has to be neither true nor false before *t*.

Kotarbiński argued that in the case if we admit indefinite propositions, that is, neither true nor false, we should revise the principle of the excluded middle. Its universally valid form is

(4) For every *A*, *A* is definite (true or false) or indefinite.

Consequently, (4) leads to rejecting

(5) True = not false.

These assertions motivate us to revise the traditional form of the principle of the excluded middle to the formula:

(6) For every *A*, *A* is true or false,

Stanisław Leśniewski (Leśniewski 1913, Leśniewski 1913a) offered two arguments against Kotarbiński. The first argues against the existence of indefinite propositions and assumes (Leśniewski as a nominalist, was speaking about sentences, not propositions) that the concept of truth is explained in the following manner:

(7) *A* is true if and only if the object signified by the subject of this sentence has the property signified by its predicate.

Briefly, a sentence is true, if it possesses the function of symbolizing (roughly speaking, this function consists in referring to something). Leśniewski argued that if the subject-term of *A* is empty, this sentence does not symbolize anything, and is false. Consequently, it may happen that a sentence *A* and its negation are false, given that the subject of *A* is empty. Thus, the principle of the excluded middle is not universally valid, and arguing against it does not require an appeal to indefinite propositions. Kotarbiński's views on this issue is that all sentences about objects prior to their creation are false. Leśniewski's second argument intends to prove that every truth is sempiternal. Assume that *A* is a truth which is not sempiternal and is now true. This means that there was a moment *t* such that *A* was not true. This entails that not-*A* was true at *t*. If we accept the principle of contradiction (no *A* is true and false), not-*A* is false now. Using the principle of contradiction once again, not-*A* is always false, which implies that it was true at *t*. Since we obtain a contradiction, every truth is sempiternal.

Leśniewski's second argument requires a definite interpretation of tensed sentences and their logical values. On this view, the temporal index *t* should be placed in the subject of the sentence. For example, the sentences 'Caesar will cross the Rubicon in 49 BC' and 'Caesar crossed the Rubicon in 49 BC' should be converted into 'Caesar-in-49-BC will cross the Rubicon' and 'Caesar-in-49-BC crossed the Rubicon', respectively. Both sentences are timeless and could be uttered in any time without worrying about their truth-values in 49 BC, earlier or later. This assumption allowed Leśniewski to consider the sentences 'Caesar will not cross the Rubicon in 49 BC' and 'Caesar crossed the Rubicon in 49 BC', uttered prior to crossing (or not) Rubicon by Caesar as mutually contradictory. For Kotarbiński, the former was indefinite, but the latter—definite. The same analysis applies to sentences about the future uttered now (at the present time). Leśniewski followed Twardowski's view that every truth is absolute. As far as the temporal aspect of truth is concerned, we can formulate the following claim:

(**AbsTr**) Truth-absoluteness = truth-eternality plus truth-sempiternality.

This definition will be somehow modified in Chap. 9, Sect. 9.5, where I will argue that **STT** defines truth as absolute.

3.7.3 Łukasiewicz

In his Łukasiewicz 1910, p. 50, he defines truth as follows:

(8) A proposition '*a* is *P*' attributing a property *P* to an object *a* is true if and only if *a* possesses *P*.

Łukasiewicz considered this formulation as well corresponding with Aristotle's intuitions. On the other hand, Łukasiewicz's interpretation of the Stagirite is fairly Brentanian, because (see p. 14), sentences are true or false provided that they state or assume that something exists or does not exist.

What about the absoluteness of truth in Łukasiewicz? He accepted this property in the sense (**AbsTr**) in 1910, but he changed his views after discovering many-valued logic. Roughly speaking, he rejected sempiternality as a mark of absolute truth. Without entering into technical details (see Chap. 4, Sect. 4.6.3), I will explain the issue with respect to three–valued logic (see Łukasiewicz 1930). Let *A* be a sentence about a future contingent event. So *A* is neither true nor false, it has the neutral (third) value, as does its negation not-*A*. Let Aristotle's example 'Sea–battle will be tomorrow' serve as an illustration (Łukasiewicz's concern was also to find a proper interpretation of the Stagirite's views about future–contingents). Thus, this sentence will be true or false tomorrow as well as for ever. This means that the absoluteness of truth is reducible to eternality. This is confirmed by a passage from Łukasiewicz 1957, p. 208:

> If truth consists in the conformity of thought to reality, we may say that those propositions are true today which conform with today's reality or with future reality in so far as that is predetermined by causes existing today. As the sae–fight of tomorrow is not real today, and its future existence has no real cause today, the proposition 'There will be a sea–fight tomorrow' is neither true not false.

Thus, Łukasiewicz accepted the correspondence definition of truth, but considered truth (and falsity) as time–dependent. Yet Łukasiewicz (see Łukasiewicz 1912) criticized the treatment of the correspondence relation as copying reality.

Łukasiewicz (still before his many-valued logic period) worked on the logical foundations of probability (see Łukasiewicz 1913). According to him, probability can be ascribed to propositional functions, that is, formulas of the for *Px*, where *x* is a free variable, but sentences (formulas without free variables) are either true or false. This view could anticipate Tarski's definition of truth as a special kind of satisfaction relation (see Chaps. 7–8). Another interesting anticipation to be found in Łukasiewicz 1913 consisted in formulating the Liar Antinomy (see Łukasiewicz 1915). As far as the principle of bivalence is concerned, Łukasiewicz, as a many-valued logician, regarded it (see Łukasiewicz 1962) as metalogical and irreducible to a concrete logical rule—such as '*A* or not-*A*', for instance. Thus, if **STT** is formulated in the metalanguage (see Chaps. 7–8), its relation to bivalence is essential.

(**DG5**) How was related Łukasiewicz's discovery of many-valued logic to the Kotarbiński–Leśniewski debate in 1913? We have no direct evidence to answer this question. See Woleński 1990a for some remarks about possible filiations.►

3.7.4 Czeżowski

According to Tadeusz Czeżowski (see Czeżowski 1919, p. 7) the sentences *A* and '*A* is true' are equivalent (an anticipation of **T**-scheme). It is perhaps interesting that Husserl (see Husserl 2002, p. 112; this is his lecture course in 1905) proposed a special case of this. Husserl and Czeżowski seem to be first philosophers who explicitly formulated **T**-scheme.

3.7.5 Later Leśniewski

In the work he published from 1920 to 1939, Leśniewski did not formulate an explicit truth-definition. Perhaps his intentions can by captured in (see Leśniewski 1931):

(9) A sentence of the form '*a* is *b*' is true if and only if '*a*' is singular non-empty term and its reference falls under '*b*'.

This definition is related to Leśniewski's ontology (calculus of names) and, for this reason, has no particular importance outside that logical system. Lęsniewski (see Leśniewski 1929) constructed a system of propositional calculus, called 'protothetic'. Since this system admits quantifying over propositional variables, its expressive power is greater than the usual sentential logic (see Chap. 5 about this latter system). In particular, we can define '*p* is (logically) false' as $\forall pp$ and '*p* is (logically) true' as $\forall p(p \Leftrightarrow p)$. More formally, we have the formulas $\mathrm{Ver}(p) \Leftrightarrow (p \Leftrightarrow p)$ and $\mathrm{Fals}(p) \Leftrightarrow (p \Leftrightarrow \neg p)$ as definitions of logical truth and falsehood; the 'official' definitions in protothetic have the form of equivalences. Consequently, the symbols Ver and Fals can be interpreted as referring to logical truth and logical falsehood. Thus, prothotetic, contrary to the standard propositional calculus enables us to define some semantic metalogical concepts. However, this results pertains to protothetic (or other systems with propositional quantifiers only. It is worthy to note that Leśniewski's ontology does not suffice to define the concept of truth as a general semantic notion. Thus, ontology does overcome Tarski's undefinability theorem (see Chap. 8). On the other hand, some other Leśniewski's ideas were influential, particularly in Poland. I mean: the sharp distinction of language and metalanguge, the diagnosis of the Liar Paradox (see Betti 2004) and so-called intuitionistic formalism. This view (having nothing in common with intuitionism in the foundations of mathematics) considers expression of any language, even formalized, as meaningful.

3.7.6 Later Kotarbiński

The next quotations are essential ((a) Kotarbiński 1926, p. 122; (b) Kotarbiński 1929, p. 106/107):

> (a) […] truth that p (the thought that p is true, the sentence 'p' is true, etc. or synonymously: the thought, that p agrees with reality) $\equiv p$.
>
> (b) Let us […] pass to the classical doctrine ask what is understood by "accordance with reality". The point is not that a true thought should be a copy or simile of the thing of which we are thinking, as a painted copy or a photograph is. A brief reflection suffices to recognize the metaphorical nature of such comparison. A different interpretation of "accordance with reality" is required. We shall confine ourselves to the following: "John thinks truly if and only if John thinks that things are so and, things are in fact so and so".

The adverbial theory is echoed in (b) (see Pasquerella 1989), but, like in Brentano (see Sect. 3.5), it results from reism as a general ontology (in Kotabiński's case, reism claims that only corporeal things exist). Both (a) and (b) propose an interpretation of the *adequatio*-formula by the equivalence of A and 'A is true'. Another view of Kotarbiński view consisted in distinguishing the verbal and real sense of 'is true'. If one say that it is true that Warsaw is the capital of Poland, the prefix 'it is true that' can be dropped, because it is enough to say that Warsaw is the capital of Poland. On the other hand, 'is true' cannot be eliminated from the context 'the theory of relativity is true'. The former usage is verbal (the nihilistic theory of truth), but the latter—real (see also (**DG6)** in Sect. 3.8).

3.7.7 Ajdukiewicz

Ajdukiewicz (Ajdukiewicz 1949, p. 18) summarized various problems connected with the correspondence theory of truth (see Ajdukiewicz 1949, p. 9):

> What is truth? The classical answer to this question states that truth of a thought consists in its agreement with reality. *Veritas est adequatiorei et intellectus*: this was classical answer in its scholastic formulation. But what is this agreement of thought and reality, as the basis of the definition of truth? Certainly not that the thought is identical with the reality it describes. Perhaps then in this, that this thought is a likeness of something real, is a reflection of reality. But even this interpretation of the 'agreement of thought and reality' seems to some philosophers an absurd idea. How, they ask, could thought be a likeness of something quite different from it, how can thought which is something that has time-dimensions but no others, be a likeness of something that is spatial.

Ajdukiewicz proposed (p. 18) the following formulation of the classical theory of truth:

> the thought *T* is true – this means: the thought *T* asserts that such-and-such is the case and such-and-such really is the case,

He considered this definition as being free of the mentioned difficulties concerning the concept of correspondence.

3.7.8 Summary

Generally speaking (I omit here some special issues, for instance, the language/metalanguage distinction; see Chap. 7, Sect. 7.4 for a discussion of this topic, very important in Tarski and other Polish philosophers and logicians, basic features of Polish (or most Polish philosophers working in the interwar period) thinking about the concept of truth, particularly as represented by the leading members of the Lvov–Warsaw School, can be summarized by the following points:

(i) the correspondence (agreement, conformity, etc.) relation as applied to the concept of truth means, to employ Kotarbiński's way of speaking, that things are as the sentence in question says they are;
(ii) the account provided in the preceding point can be called classical and closely related to Aristotle's ideas;
(iii) truth is absolute (eternal and sempiternal);
(iv) bearers of truth are judgments, propositions, sentences, etc.—let us say they items of the propositional category (in the syntactic sense; this admits to use terms 'sentence' and 'proposition' as equivalent in syntactic contexts).

Polish philosophers did not avoid (they did not try to do that) terms like 'correspondence', 'agreement', 'conformity', etc., and they understood or even defined them in a way regarded as precise. In order to have convenient labels, I will distinguish (see Woleński, Simons 1989, p. 399, Woleński 1993) the strong correspondence (there is variety of related accounts, but that of Russell is a good example; see **DG2**) and weak correspondence (as in Twardowski and his followers). Historically speaking, the link between the Polish tradition and Brentano's views is evident.

3.8 Some Truth-Definitions in the 20th and 21st Century

This section presents a sample of truth-definitions or descriptions of the concept of truth (not all quoted passages are definions in the strict sense) from the 20th century and the beginning of 21st century. I will not make extensive comments about them, because it would be exceed the scope of this book. Here is the list:

(i) Husserl 1913, v. 2, p. 263 (see also Sect. 3.7.4):

> *Truth* […] [is] *the full agreement* of what is meant *with what is give* [a state of affairs] *as such.*

(ii) Schlick 1918, p. 61

> A judgement that *uniquely designates* a set of facts is called *true* […] the concept of truth was almost always defined as an agreement between thought and its object – or better, between judgement and what is judged. […] this definition expresses a correct conception. […] the notion of agreement, in so far as it is to mean sameness or similarity, melts away under the rays of analysis, and what is left is unique coordination. It is the latter that the relationship of true judgements consists, and all those naive theories according to which our judgements and concepts are able in some fashion to "picture" reality are completely demolished. No other sense remains for the word "agreement" that that of unique coordination or correspondence.

(iii) Wittgenstein 1922:

> 4.011 A proposition is a picture of reality. […] A proposition is a model of reality.
> 4.022 […] A proposition *shows* how things stand *if* it is true.
> 4.05 Reality is compared with proposition.
> 4.06 Propositions can be true or false only by being pictures of reality.

(iv) Ramsey 1927, p. 143:

> The propositional function p is true is simply the same as p.

(v) Ramsey 1991, p. 9 (in fact this definition was formulated in 1927–1929):

> a belief is true if and only if it is a belief that p and p.

(vi) Ayer 1946, p. 117/118:

> Reverting to the analysis of truth, we find that in all sentences of the form 'p is true', the phrase 'it true' is logically superfluous. When, for example, one says that that the proposition 'Queen Anne is dead' is true, all that one is saying is that Queen one is dead. Thus, to say that a proposition is true is just to assert it, and to say that that it is false is just to assert its contradictory. And this indicates that the terms 'true' and 'false' connote nothing, but its function in the sentence simply as marks of assertion and denial.

(vii) Carnap 1947, p. 5:

> An atomic sentence […] consisting of a predicate followed by an individual constant is true if and only if the individual to which the individual constant refers possesses the property to which the predicate refers.

(viii) Popper 1972, p. 44:

> I accept the commonsense theory (defended and refined by Alfred Tarski) that truth is correspondence with facts (or wb ith reality); or more precisely, that a theory is true if and only if it corresponds to the facts.

(ix) Quine 1987, p. 213:

> The combination 'it is a fact that' is vacuous […]. 'It is a fact that snow is white' reduces to 'Snow is white'. Our account of the truth of 'Snow is white' in terms of facts has now come down to this: 'Snow is white' if and only if snow is white. […] Here, as Tarski, has urged, is the significant residue of the correspondence theory of truth. To attribute truth to the sentence is to attribute whiteness to snow. Attribution of truth to 'Snow is white' just cancels the quotation marks and says that snow is white. Truth is disquotation.

(x) Grover 1992, pp. 88–89:

> 'That is true' and 'It is true' can be and should be thought as anaphoric prosentences […] For each proposition, if John said that it is true, then it is true.

(xi) Alston 1996, p. 5:

> A sentence (proposition, belief …) is true if and only if what the statement says to be the case actually is the case.

(xii) Horwich 1998, p. 6:

> It is true that *p* if and only if *p*.

(xiii) Dodd 2000, p. 111:

> A fact is a thought that is true. […] <*p*> is true if and only if <*p*> is *identical* with a fact.

(xiv) Merricks 2007, p. 170:

> *Being true* is a primitive monadic property.

(xv) Frápolli 2013, p. 11:

> Truth is a higher-order concept thanot represent any trait of the external world.

(xvi) Czarnocka 2017, p. 187

> *The correspondence of reality with knowledge consists in symbolizing*; the correspondence relation connect the cognitive objerct with its symbols.

I intentionally included Husserl in order to give an example of the *adequatio*-formula as a component of a very complex aletheiological theory based on premises rather remote from other proposals mentioned in this section. It is an interesting circumstance, because documents that intuitions behind the concept of correspondence are (partially, in order to be careful) independent of mutually very different philosophical background. Moritz Schlick's approach was semantic for being based on the concept of designation attributed to judgments (or propositions). He also criticized strong correspondence and, in fact, Ajdukiewicz repeated his arguments. Carnap's definition is semantic. The concept of strong correspondence relation occurs in (iii), (viii) and (xvi). Other listed accounts (perhaps except (xiv)) fall under the minimalist (deflanionist, redundantism, disquotationism) approach which maintains that the expression 'it is true that' can be eliminated, due to the equivalence of *A* and '*A* is true'. Since this scheme is central in **STT**, it is quite natural to compare minimalist truth theories with that of Tarski. In fact, some minimalists claim that they simplify **STT**, but achieve its goals. I will return to this issue in Chap. 9, Sect. 9.7. As I already noted, I abstain from commenting on (i)–(xii). This move simplifies my description. It is perhaps the most evident with the respect to (xii) (the identity theory) and (x) (the prosentential theory), which are much more sophisticated in their original versions. In particular, I do not analyze the concepts of prosentence (it is interesting that this concept was introduced in Brentano 1930, p. 65, but not for analyzing truth), and identity of <*p*> (the symbol <···> is another device to indicate that the expression inside it is mentioned, not used; see also Chap. 7, Sect. 7.4), and fact. Anyway, aletheiological minimalism is decisively based on the weak notion of correspondence.

(**DG6**) Frank Ramsey discussed the problem whether (iv) is sufficiently general (see Ramsey 1927, p. 143). Consider the sentence (*) 'He is always right' as equivalent to (**) 'for all *p*, if he asserts *p*, *p* is true'. For the first look, 'true' cannot be eliminated from (**). Ramsey proposes (***) 'for all *a*, *R*, *b*, if he assert *aRb*, then *aRb*' and says about it "to which 'is true' would be an superfluous addition". Ramsey considered this question already in 1922 (thanks to Michael Potter for information), in a talk delivered to the Cambridge Apostles (see Ramsey 2007). He remarked: "I say that p is true is merely a different verbal form for p. If however we consider 'He's said something true' we cannot dispose of the matter as easily as this". Let me remind (see F above) that Kotarbiński argued that 'is true' is not eliminable from such contexts (see also Chap. 9, Sect. 9.7 and probably would agree with Ramsey about "He's said something true'.►

(**DG7**) At this place, some remarks on **T**-scheme are required. Consider a concrete example (such concrete exemplifications are called **T**-equivelences or **T**-sentences—from the logical point of view they can be viewed as particularizations or substitutions of **T**-scheme)

(a) 'Warsaw is the capital of Poland' is true if and only if Warsaw is the capital of Poland.

A more complex version of (a) is

(b) the sentence 'Warsaw is the capital of Poland' is true if and only if Warsaw is the capital of Poland.

Clearly, we can drop 'the sentence' in (b) and stay with (a). The question is how (a) is related to the formula

(c) *A* is true if and only if *A*.

The quotes in (a) indicate that the quoted sentence is mentioned, but this sentence (without quotes) is used in the right part of (a) (this statement records the famous use/mention distinction). Now, the difference between use and mention has no indication in (c). The letter *A* functions as a metavariable, that is a representation of any sentence (in the sense of logic; see Chap. 5 for a closer characterization. Consequently, (c) means (d) a sentence represented by the metavariable *A* is true if and only if *A*. Under this convention, (a) is a particularization of (d). Other way to introduce the use/mention distinction into (c) consists in writing (e) '*A*' is true if and only if *A*. The simplest interpretation of the expression '*A*' (the metavariable *A* in quotes) consist in saying that it is a name of a sentence represented by *A*. If these explanations are taken into account, we can use (c) in informal considerations. This version is frequently termed as the naive **T**-scheme. Tarski showed why the naive version should be modified in order to function in frameworks of a satisfactory truth-definition, in particular with respect to semantic paradoxes (see Chap. 7).

(**DG8**) I do not enter into various problems related to the concept of correspondence. For instance, Wolfgang Künne (see Künne 2003, pp. 3–5 and David 2004, pp. 338–343) considers the division of objects-based correspondence theories and facts-based correspondence theories as particularly important. The former sees the essence of truth in relation of what is true to object (Brentano's approach can be taken as a typical example), but the latter defines truth by relation to facts (Russell's theory provides an illustration). Since an analysis of **STT** does not require a reference to Künne's distinction I omit it from my further considerations. In particular, I regard the distinction of weak-correspondence theories and strong-correspondence theories as more significant in the context of **STT** than the classification introduced by Künne (in fact, his map of truth-theories is much complex than the mentioned opposition). Yet I agree that a detailed historical presentation of various aletheiologies would profit from taking into account the difference between objects-oriented and facts-oriented correspondence theories. See also Hoven 1989 for various typologies of truth-theories; see also Chap. 4, Sect. 4.4.►

(**DG9**) This chapter is partially based on Woleński, Simons 1989, Woleński 1994b, Woleński 1998, Woleński 2004, Woleński 2009, Woleński 2015 and Murawski, Woleński 2008.►

References

Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E., Romanowska, A., Woleński, J. (Eds.) (2004). *Provinces of Logic Determined. Essays in the Memory of Alfred Tarski. Parts I, II and III* (*Annals of Pure and Applied Logic*, 127). Amsterdam: Elsevier.

Ajdukiewicz, K. (1949). *Zagadnienia i kierunki filozofii*. Warszawa: Czytelnik; Eng. tr. (by H. Skolimowski, A. Quinton), *Problems & Theories of Philosophy*. Cambridge: Cambridge University Press.

Alston, W. (1996). *A Realist Conception of Truth*. Ithaca: Cornell University Press.

Ayer, A. J. (1946). *Language, Truth and Logic*. London: Gollancz (2nd ed.).

Betti, A. (2004). Leśniewski's aarly liar, Tarski and natural language. In Adamowicz, Artemov, Niwiński, Orłowska, Romanowska, Woleński (2004), 267–287.

Bolzano, B. (1837). *Wissenschaftslehre*. Sulzbach: Geiblish; Eng. tr., *Theory of Science*. Oxford: Oxford University Press 2014.

Bosanquet, B. (1911). *Logic or the Morphology of Knowledge.* Oxford: Clarendon Press.

Bradley, F. H. (1883). *The Principles of Logic*. Oxford: Clarendon Press.

Bradley, F. H. (1893). *Appearance and Reality. A Metaphysical Essay*. Oxford: Clarendon Press.

Bradley, F. H. (1907). On truth and copying. Mind, XVI, 94–110; repr. in Bradley (1914), 107–116.

Bradley, F. H. (1914). *Essays on Truth and Reality*. Oxford: Clarendon Press.

Bradley, F. H. (1922). *The Principles of Logic.* Oxford: Clarendon Press (2nd ed.).

Brandl, J. (2017). Was Brentano an early deflationist about truth? *The Monist*, 100(1), 1–14.

Brentano, F. (1930). *Wahrheit und Evidenz*. Leipzig: Felix Meiner; Eng. tr. (by R. Chisholm, I. Politzer, K. Fischer), *The True and the Evident*. London: Routledge and Kegan Paul 1966.

Brożek, A., Chybińska, A., Jadacki, J., Woleński, J. (Eds.) (2015). *Tradition of the Lvov–Warsaw School*. *Ideas and Continuations.* Leiden: Brill/Rodopi.

Candlish, S. (2007). *The Russell/Bradley Dispute and Its Significance for Twentieth Century Philosophy*. Basingstoke: Palgrave Macmillan.

Carnap, R. (1947). *Meaning and Necessity. A Study in Semantics of Modal Logic*. Chicago: University of Chicago Press; 2nd ed., 1956.

Casari, E. (2016). *Bolzano's Logical System*. Oxford: Oxford University Press.

Church, A. (1943). Carnap's *Introduction to Semantics*. *The Philosophical Review*, 52, 298–305; repr. in Church (2019), 259–264.

Church, A. (2019). *Collected Papers.* The MIT Press.

Czarnocka, M. (2017). *A Path to a Conception of Symbolic Truth.* Frankfurt am Main: Peter Lang.

Czeżowski, T. (1919). Imiona i zdania (Names and sentences). *Przegląd Filozoficzny*, 22, 1–19.

David, M. (2004). Theories of truth. In Niiniluoto, Sintonen, Woleński (2004), 331–414.

Davidson, D. (1969). True to facts. *The Journal of Philosophy*, 66, 748–764; repr. in Davidson (1984), 37–54.

Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.

Dodd, J. (2000). *The Identity Theory of Truth*. London: Macmillan.

Eisler, R. (1930). *Wörterbuch der philosophischen Begriffe*, v. III. Berlin: Mittler.

Erdmann, B. (1892). *Logik. Logische Elementarlehre*. Halle: Niemeyer; Berlin: de Gruyter (3rd ed.).

Frápolli, M. J. (2013). *The Nature of Truth: An Updated Approach to the Meaning of Truth Ascriptions*. Berlin: Springer.

Frege, G. (1918). Der Gedanke. *Beiträge zur Philosohie des deutschen Idealismus*, I, 58–77; Eng. tr. (by P. Geach). In Frege (1984), 351–372.

Frege, G. (1969). *Nachgelassene Schriften*. Hamburg: Felix Meiner; Eng. tr., *Posthumous Writings*. Oxford: Basil Blackwel 1979.

Frege, G. (1984). *Collected Papers on Mathematics, Logic and Philosophy*. Oxford: Blackwell.

Fries, J. F. (1837). *System der Logik. Ein Handbuch zur Lehrer und Selbstgebrauch*. Heidelberg: Winter.

Gödel, K. (1944). Russell's mathematical logic. In Schillp (1944), 125–153; repr. in Gödel (1989), 119–141.

Gödel, K. (1989). *Collected Works*, v. II. Oxford: Oxford University Press.

Gotthardt, J. (1918). *Das Wahrheitsproblem in dem philosophischen Lebenswerk Bernard Bolzanos*. Trier: Paulinus-Drückeri.

Greimann, D. (Ed.) (2007). *Essays on Frege's Conception of Truth* (*Grazer Philosophische Studien*, 75). Amsterdam: Rodopi.

Grover, D. (1992). *A Prosentential Theory of Truth*. Princeton: Princeton University Press.

Hintikka, J., Puhl, K. (Eds.) (1994). *The British Tradition in 20th Century Philosophy Papers of 17th Ludwig Wittgenstein Symposium*. Kirchberg am Wechsel: The Austrian Ludwig Wittgenstein Society.

Horwich, P. (1998). *Truth.* Oxford: Clarendon Press; 2nd ed., 2000.

Hoven, A. (1989). *Wege zur Wahrheit. Eine typologische Studie über Wahtheitstheorien.* Frankfurt am Main: Peter Lang.

Husserl, E. (1913). *Logische Unresuchungen* (Band 2). Halle: Niemeyer; Eng. tr. (by J. N. Findlay), *Logical Investigations* (v. 2). London: Routledge 2001.

Husserl, E. (2002). *Urteilstheorie. Vorlesung 1905*. Dordrecht: Kluwer.

James, W. (1907). *Pragmatism. A New Name for Some Old Ways of Thinking. Popular Lectures on Philosophy*. London: Longmans, Green and Co.

Joachim, H. (1906). *The Nature of Truth. An Essay.* Oxford: Clarendon Press.

Joachim, H. (1907). A reply to Mr. Moore. *Mind*, XVI, 410–415.

Kotarbiński, T. (1913). Zagadnienie istnienia przyszłości (The problem of the existence of future). *Przegląd Filozoficzny*, 15, 74–92; Eng. tr, *The Polish Review*, 13(3)(1968), 7–22.

Kotarbiński, T. (1926). *Elementy logiki formalnej, teorii poznania i metodologji* (Elements of Formal logic, Theory of Knowledge and Methodology). Warszawa: Wydawnictwo Koła Filozoficznego S. U. W. i Koła Przyrodników S. U. W.

Kotarbiński, T. (1929). *Elementy teorii poznania, logiki formalnej i metodologii nauk* (Elements of Theory of Knowledge, Formal Logic and Methodology of Science). Lwów: Ossolineum; Eng. *Gniosiology. The Scientific Approach to the Theory of Knowledge*. Oxford: Pergamon Press.

Künne, W. (2003). *Conceptions of Truth.* Oxford: Clarendon Press.

Lapointe, S., Woleński, J., Marion, M., Miśkiewicz, W. (2009). *The Golden Age of Polish Philosophy. Kazimierz Twardowski's Philosophical Legacy*. Berlin: Springer.

Leśniewski, S. (1913). Krytyka logicznej zasady wyłączonego środka (The critique of the logical principle of the excluded middle). *Przegląd Filozoficzny*, 16, 315–352; Eng. tr. in Leśniewski (1992), 17–85.

Leśniewski, S. (1913a). Czy prawda jest tylko wieczna czy też wieczna i odwieczna (Is all truth only true eternally or it is also true without a beginning). *Nowe Tory*, 10, 493–528; Eng. tr. in Leśniewski (1992), 86–114.

Leśniewski, S. (1929). Grundzüge eines neuen System der Grundlagen der Mathematik. *Fundamenta Mathematicae*, 14, 1–81; Eng. tr. (by M. P. O'Neil) In Leśniewski (1992), 410–605.

Leśniewski, S. (1931). O podstawach matematyki VI (On the foundations of mathematics). Przegląd Filozoficzny, 34, 142–170; Eng. tr. in Leśniewski (1992), 364–382.

Leśniewski, S. (1992). *Collected Works*. Dordrecht: Kluwer.

Łukasiewicz, J. (1910). *O zasadzie sprzeczności u Arystotelesa. Studyum Krytyczne* (On the Principle of Contradiction in Aristotle). Kraków: Wydawnictwo Polskiej Akademii Umiejętności; Eng. tr (by H. R. Heine), *The Principle of Contradiction in Aristotle: A Critical Study*. North Charleston, SC: CreateSpace Independent Publishing 2018.

Łukasiewicz, J. (1912). O twórczości w nauce (Creative elements in science). In *Księga Pamiątkowo ku uczczeniu 250-tej rocznicy założenia Uniwersytetu Lwowskiego przez króla Jana Kazimierza* (Memorial Book on the Occasion of Establishment of Lvov University by the King Jan Kazimierz), v. II, 1–20. Lwów: Nakładem Uniwersytetu Lwowskiego, 1–20; Eng. tr. in Łukasiewicz, 1–15.

Łukasiewicz, J. (1913). *Die logische Grundlagen der Wahrshenlichkeitsrechnung*. Kraków: Wydawnictwo Polskiej Akademii Umiejętności; partial Eng. tr. in Łukasiewicz (1970), 16–63.
Łukasiewicz, J. (1915). O nauce (On science). In *Poradnik dla Samouków*, Warszawa: Heflich i Michalski, XV–XXXIX).
Łukasiewicz, J. (1930). Philosophische Bemerkungen zu mehrwertigen Systemem des Aussagenkalküls. *Comptes Rendus des Sèances de la Société des Sciences et de Lettres de Varsovie, Classe III*, 51–77; Eng. tr. in Łukasiewicz (1970), 153–178.
Łukasiewicz, J. (1957). *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic*. Oxford: Clarendon Press.
Łukasiewicz, J. (1962). O determinizmie [Mowa Rektorska z 1922 r.] (On determinism. The Rector Speech in 1922]. In Łukasiewicz (1962), 114–126; Eng. tr. in Łukasiewicz (1970, 110–128.
Łukasiewicz, J. (1962). Z zagadnień logiki i filozofii (Problems of Logic and Philosophy). Warszawa: Pańśtwowe: Wydawnictwo Naukowe.
Łukasiewicz, J. (1970). *Selected Works*. Amsterdam: North-Holland.
Mauthner, F. (1902). *Zur Grammatik und Logik*. Stuttgart: Cotta.
Meier, H. (1926). *Wahrheit und Wirklichkeit*. Tübingen: Mohr.
Merricks, T. (2007). *Truth and Ontology*. Oxford: Clarendon Press.
Moore, G. E. (1907). Mr. Joachim's nature of truth. *Mind*, XV, 229–235.
Moore, G. E. (1953). *Some Main Problems of Philosophy*. London: George Allen & Unwin.
Murawski, R., Woleński, J. (2008). Tarski and his precedessors on truth. In Patterson (2008), 21–43.
Neale, S. (2001). *Facing Facts*. Oxford: Clarendon Press.
Nesher, D. (2002). *On Truth and the Representation of Reality. A Collection of Inquiries from a Pragmatist Point of View*. Landham: University Press of America.
Niiniluoto, I., Sintonen, M., Woleński, J. (Eds.) (2004). *Handbook of Epistemology*. Dodrecht: Kluwer.
Pardey, U. (2012). *Frege on Absolute and Relative Truth*. New York: Palgrave Macmillan.
Pasquerella, L. (1989). Brentano and Kotarbiński on truth. *Topoi*, 4 (Supplement: *Object and Its Identity*), 98–106.
Patterson, D. E. (Ed.) (2008). *New Essays on Tarski and Philosophy*. Oxford: Oxford University Press.
Popper, K. (1972). *Objective Knowledge. An Evolutionary Approach*. Oxford: Clarendon Press; repr. with additions, 1974.
Prawitz, D., Skyrms, B., Westerståhl, D. (Eds.) (1994). *Logic, Methodology and Philosophy of Science IX*. Amsterdam: Elsevier
Quine, W. v. O. (1987). *Quiddities. An Intermittently Philosophical Dictionary*. Cambridge, Mass.: Harvard University Press.
Ramsey, F. P. (1927). Facts and propositions. *Proceeding of Aristotelian Society. Supplementary Volume*, VII, 153–170; repr. in Ramsey (1931), 138–155 and Ramsey (1978), 40–57.
Ramsey, F. P. (1931). *The Foundations of Mathematics*. London: Routledge and Kegan Paul; 2nd ed., *Foundations*. *Essays in Philosophy*, *Logic*, *Mathematics and Economics*. London: Routledge and Kegan Paul 1978.
Ramsey, F. P. (1991). *On truth. Original Manuscript Materials (1927–1929) from the Ramsey Collection at the University of Pittsburgh*. Dordrecht: Kluwer.
Ramsey, F. P. (2007). Truth and simplicity (Ed. by A. Koslow). *British Journal of Philosophy of Science*, 58, 1–7.
Russell, B. (1903). *The Principles of Mathematics*. Cambridge: Cambridge University Press.
Russell, B. (1904). Meinong's theory of complexes and assumptions (III). *Mind*, XIII, 509–524; repr. in Russell (1973), 54–69.
Russell, B. (1906). The nature of truth. *Mind*, XV, 528–533.
Russell, B. (1906a). What is truth? *Independent Review*, 9, 349–353.
Russell, B. (1907). On the nature of truth. *Proceedings of Aristotelian Society (New Series)*, VII, 28–49.

Russell, B. (1910). *Philosophical Essays*. London: Longmans.

Russell, B. (1912). *The Problems of Philosophy*. London: Wiliam and Norgate.

Russell, B. (1940). *An Inquiry into Meaning and Truth.* London: George Unwin and Allen.

Russell, B. (1973). *Essays in Analysis*. New York: George Braziller.

Russell, B. (1984). *The Theory of Knowledge. The 1913 Manuscript*. London: George Allen and Unwin.

Schillp, P. A. (Ed.) (1944). *The Philosophy of Bertrand Russell*. La Salle: Open Court.

Schlick, M. (1918). *Allgemeine Erkenntnislehre*. Springer, Berlin; Eng. tr. *General Theory of Knowledge*. Wien: Springer.

Simons, P. (1992). *Philosophy and Logic in Central Europe from Bolzano to Tarski. Selected Essays.* The Hague: Nijhoff.

Srzednicki, J. (1965). *Franz Brentano's Analysis of Truth*. The Hague: Nijhoff.

Tarski, A. (1992). Drei Briefe an Otto Neurath [25. IV. 1930, 10. VI. 1936, 7. IX. 1936], together with Eng. tr., *Grazer Philosophische Studien*, 42, 1–29.

Twardowski, K. (1900). O tzw. prawdach względnych (On so-called relative truths). In *Księga Pamiątkowa Uniweysytetu Lwowskiego ku uczczeniu pięćsetnej rocznicy Fundacji Jagiellońskiej Uniwersytetu Krakowskiego* (Memorial Book of Lvov University on the Occasion of 500th Anniversary of Jagiellonian Fund—the University of Kraków) Lwów: Nakładem Uniwersytetu Lwowskiego, 64–93; Eng. tr. in Twardowski (1999), 103–122.

Twardowski, K. (1912). O czynnościach i wytworach. Kilka uwag z pogranicza psychologii, gramatyki i logiki (Actions and products. Remarks from the boderline of psychology, grammar and logic). In *Księga Pamiątkowo ku uczczeniu 250-tej rocznicy założenia Uniwersytetu Lwowskiego przez króla Jana Kazimierza* (Memorial Book on the Occasion of Establishment of Lvov University by the King Jan Kazimierz), v. II. Lwów: Nakładem Uniwersytetu Lwowskiego, 1–33; Eng. tr. in Twardowski (1999), 103–122.

Twardowski, K. (1975). Wykłady z teorii poznania w r. a. 1924–25 (Theory of Knowledge. Lecture Course 1924–25). *Archiwum Historii Filozofii i Myśli Społecznej*, 22, 241–299; Eng. tr. in Twardowski (1999), 181–239.

Twardowski, K. (1999). *On Actions, Products and Other Topics in Philosophy*. Amsterdam: Rodopi.

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. London: Routledge and Kegan Paul.

Woleński, J. (Ed.) (1990). *Kotarbiński: Logic, Semantics and Ontology*. Dordrecht: Kluwer

Woleński, J. (1990a). Kotarbiński, many-valued logic and truth. In Woleński (1990), 190–197; repr. in Woleński (1999), 115–120.

Woleński, J. (1993). Two concepts of correspondence. *From the Logical Point of View*, II(3), 42–55.

Woleński, J. (1994). Controversy over the concept of correspondence (Bradley, Joachim, Russell). In Hintikka, Puhl (1994), 537–543; repr. in Woleński (1999), 233–236.

Woleński, J. (1994b). Contributions to the history of the classical truth-definition. In Prawitz, Skyrms, Westerståhl (1994), 481–495; repr. in Woleński (1999), 139–149.

Woleński, J. (1998). Theories of truth in Austrian philosophy. *Reports on Philosophy*, 18, 13–49; repr. in Woleński (1999), 150–175.

Woleński, J. (1999). *Essays in the History of Logic and Logical Philosophy*. Kraków: Jagiellonian University Press.

Woleński, J. (2004). Aletheia in Greek thought until Aristotle. In Adamowicz, Artemov, Niwiński, Orłowska, Romanowska, Woleński (2004), 339–360.

Woleński, J. (2009). The rise and development of logical semantics in Poland. In Lapointe, Woleński, Marion, Miśkiewicz (2009), 43–59.

Woleński, J. (2015). Truth-theories in the Lvov–Warsaw School. In Brożek, Chybińska, Jadacki, Woleński, 2015, 73–92.

Woleński, J., Simons, P. (1989). De Veritate: Austro-polish contributions to the theory of truth from Brentano to Tarski. Szaniawski, 1989, 391–442.

Chapter 4
Tasks of Truth-Theories

Abstract This chapter concentrates on problems discussed in truth theories, namely truth-bearers, the question whether the concept of truth belongs to ontology, epistemology or axiology, the definability of truth, truth-criteria, formal properties of the division of truth-bearers into truths and falsehoods, truth and logic, relations between the concept of truth and other philosophical notions, the applicability of the concept of truth to common sense, science, art, religion, morality, etc.

4.1 Introduction

Every philosophical theory of truth should answer several questions. I propose the following list of problems for a truth-theory:

(I) What are bearers of truth?
(II) Is truth an ontological, epistemological, or axiological concept?
(III) Is truth definable?
(IV) If the answer to (III) is "yes", how is truth definable?
(V) What are truth-criteria and how are they related to truth-definitions?
(VI) Is the division of truths and falsehoods exhaustive and strongly disjunctive?
(VII) Is the division into truths and falsehoods stable?
(VIII) What about the logical properties of the concept of truth?
(IX) How is the concept of truth related to other philosophical concepts, and applied to various philosophical problems?
(X) Is the concept of truth applicable to common sense, science, art, religion, law, morality, etc.?

(**DG1**) More or less extensive discussions on (I)–(X) (as well as other problems touched in this book) are to be found in every introductory treatment of aletheiology, as well as advanced monographs on this subject. I mention only eight books in English, namely Williams 1976, Johnson 1992, Kirkham 1992, Soames 1999, Weingartner 2000, Halbach, Horsten 2002, Schantz 2001, Künne 2003, Williams 2004, Edwards 2018, Jago 2018, and Puntel 1990 in German.►

J. Woleński, *Semantics and Truth*, Logic, Epistemology, and the Unity of Science 45, https://doi.org/10.1007/978-3-030-24536-8_4

(**DG2**) I do not claim that that the above list is exhaustive. I am mostly interested in questions relevant to **STT**. This orientation dictates that my report be selective. For instance, I omit altogether the problem of truth in art (see Hofstadter 1965, Gaskin 2013), literary fiction (see Woods 2018) and religion (see Vroom 1989). The question of truth in law (see Patterson 1999) has special dimension in some constraints concerning the criteria of establishing facts before courts (for example, presumptions or the requirements of direct evidence). The issue of moral truth leads to the question of logical values of normative and evaluative statements (see McCloskey 1969). A new direction of research (see Changeux 2002) tries to explain searching for truth by tools used in neuroscience. As a convinced naturalism, I consider this approach as very interesting, but it does not belong to semantics. To round out the complete omissions, I except for some parenthetical remarks, into an analysis of such contexts as 'true speech acts', 'true actions', 'true intentions', 'truthfulness', 'true love', 'true (veridical, matching reality) perception', etc. (see Halldén 1960, Bennett, Hoffman, Prakash 1989, Albuquerque 1995, Williams 2004).►

4.2 Truth-Bearers: General Remarks

(**DG3**) Such terms as 'truth-bearer' and 'truth-value' are truth-oriented are labels used in this field of philosophy. However, we can equally well speak about falsehood-bearers. The parity between truth-bearers and falsehood-bearers is displayed by the following formula:

(*) for any *A*, *A* can be a truth bearer if and only *A* can be a falsehood-bearer.

Assuming that $\neg A$ s read '*A* is false', (*) might be rephrased as the statement that *A* can be a truth-bearer if and only if $\neg A$ can be For obvious logical reasons, we cannot say that *A* is a truth-bearer if and only if *A* is a falsehood-bearer, because that would seem to suggest that an *A* could be simultaneously true and false (see Sects. 4.7 and 4.8 for further remarks about this problem). The formula (*) means that an item, which can function as a possible truth-bearer, is also a possible falsehood-bearer. Some issues concerning truth—for example, the nature of truth-bearer—have the same significance for falsehood (compare Russell's conditions for a correct truth-theory mentioned in Chap. 3, Sect. 3.3). Perhaps a more neutral terminology would be proper in such cases, for instance, employing the term 'logical value-bearer' (this proposal is consistent with admitting other logical values than truth and falsehood; yet there is a problem with so-called truth-value gaps, that is bearers without having logical values). However, I will adhere to traditional terminology because it became standard. This section is based on Woleński 2004e.►

The problem of truth-bearers has a very simple formulation. Consider the scheme:

(1) x is true.

We ask now what kind of objects (items, entities, etc.) are values of the variable x occurring in (1). Generally speaking, one can distinguish nominal and sentential (or propositional) conceptions of truth-bearers. The criterion of this distinction takes into account the categories of objects represented by the variable x. I take a linguistic approach to this problem, that is, to say, I identify the syntactic category of what can be substituted for the variable x in (1) (this explanation will be qualified later). The nominal theories are reducible to those that propose concepts (or something like them) as truth-bearers. In other words, nominal theories in the linguistic setting regard names as expressing concepts. Doctrines of Hegel, some Neo-Hegelians (but not Bradley, for instance), or James (but not all pragmatist) might serve as historical examples; according to James, truth functions as an attribute of ideas, which are comparable with presentations. Since the nominal theory of truth-bearers has no greater significance for my further considerations, I confine myself to this very general and rather simplified account.

A general feature of the sentential theories of truth-bearers consists in taking truth as attributable to sentences or entities expressed by sentences. This suggests that (1) should be extended to:

(2) What is expressed by a sentence x, is true.

This formula opens many possibilities that have been exploited in the philosophical past (see Chaps. 1, 2, and 3). The list of candidates comprise (the denominations in brackets are examples): sentences (Leśniewski), propositions (Twardowski, Pap), statements (Strawson), speech-acts (Austin), thoughts (Frege), beliefs (Russell, James), judgments (older logical literature, but also Russell) or acts (Brentano). According to this variety of options, we have the following list of various concretisations of (2):

(3) Sentences are true;
(4) Propositions expressed by sentences are true (provided that one sentence expresses one proposition);
(5) Statements expressed by sentences are true;
(6) Thoughts expressed by sentences are true;
(7) Beliefs expressed by sentences are true;
(8) Judgments expressed by sentence are true;
(9) Acts (of judging) expressed by sentences are true.

(**DG4**) View (3) requires an additional assumption, somewhat artificial but tolerable, that sentences are self-expressible; it means that is A is a sentence, it expresses itself. I do not claim that I mentioned all possible candidates for truth-bearers, because I omitted, utterances or assertions, for example. Another way to extend the class of possible truth-bearers is to attribute truth-values to guesses, hypotheses, assumptions, presumptions, etc. Philosophically that seems dubious (because one can claim, for instance, that hypotheses are neither true nor false, but only probable), but it is acceptable in common usage. See Engel 1991, Kirkham 1992, pp. 54–67, David 2004, Sect. 1 and Rojszczak 2004 for more detailed accounts of what can function as a truth-bearer.►

Consider now the sentence ‘snow is white’. As particular instances of (3)–(9) we have, respectively (I omit (9)):

(10) The sentence ‘snow is white’ is true;
(11) The proposition expressed by the sentence ‘snow is white’ is true;
(12) The statement expressed by the sentence ‘snow is white’ is true;
(13) The thought expressed by the sentence ‘snow is white’ is true;
(14) The belief expressed by the sentence ‘snow is white’ is true;
(15) The judgment expressed by the sentence ‘snow is white’ is true;

All of the examples (10)–(15) can be captured more simply by

(16) The proposition (statement, thought, belief, judgment, act of judging) that snow is white is true.

However, we cannot do the same with (10), because the context:

(17) The sentence that snow is white is true,

Does not capture what is going on, since it is not directly about the sentence ‘snow is white’ but reports rather what this sentence expresses. In fact, even a more serious objection can be formulated, namely that (17) is an example of an expression in which a category-mistake is involved.

In general, reports in *oratio oblique* contexts induce the introduction of truth-bearers other than sentences (see Dummett 1999, p. 1, referring to Frege). On the other hand, all (10)–(15) fall under:

(18) It is true that snow is white,

This shows that all the cases designated in (10)–(15) have something in common. Now, propositions are natural candidates to occur after ‘It is true that’ (this expression is a propositional operator) in (18). However, everything depends on how they are understood. One possibility consists in identifying them with judgments, but if propositions are conceived as abstract entities (for example, Fregean thoughts) or classes of possible worlds, the language of acts and products is not a proper analytic device. Likewise, (18) as applied to propositions cannot be equally well used in the case of sentences. Thus, we have two general approaches to truth-bearers. The first consists in taking sentences as truth-bearers, the second favours propositions.

The distinction of sentences and propositions seems natural, but is perplexing. The pairs of words ‘sentence’—‘proposition’ (English), *enuntiatio—iudicium* (Latin), *Aussage—Urteil* (German) or *zdanie—sąd* (Polish) attest that this distinction is common in various languages. The history of this distinction is rich and important for logic, semantics, epistemology, ontology and psychology (see a survey in Van Zantwijk, Gabriel 2001). Since I cannot deal with the issue of propositions here, in particular their ontological nature, I will restrict myself to a few very elementary questions, which have (or seem to have) a direct connection with semantics and the problem of truth-bearers. I do not give bibliographical references that pertain to particular views about propositions since these are easily accessible in the texts listed below. Fortunately, the discussion of this variety of this topic does not require a treatment of subtle and difficult ontological problems.

It can be said that semantics employs propositions in order to thwart the psychologistic explanation of the concept of meaning. To see that, consider the sentences (i) 'snow is white' and (ii) *Schnee ist weiss* (recall that expression other than English are italicized). Naturally, we will say that (i) means in English the same as (ii) in German. Both sentences are completely different from the lexical point of view, although they consist in three words. On the other hand, (i) and (ii) mean the same can be explained by saying that although these sentences are lexically different, they express the same propositions. So far so good, but the matter becomes more complicated when we ask 'What are propositions?' In a very broad sense (see Engel 1991, p. 376), propositions are items that are asserted, rejected, judged, believed, etc. Furthermore, propositions are conceived as linguistic (or at least very closely associated with language) entities suitable to assert, etc. something. Thirdly, propositions are also seen as intensional entities that constitute the meaning of sentences (that was our actual starting point). A helpful list compiled in Kirkham 1992, pp. 55–56 contains the following entries: (a) psychical entities; (b) contents of utterances; (c) meanings of sentences; (d) objects of consciousness; (e) what is common to synonymous sentences; (f) entities outside of space and time; (g) what is common to sentences in various grammatical moods (for example, 'the windows are closed' and 'are the windows closed?'); (g) facts. Perhaps this list becomes clearer if we compare it with an old distinction between propositions in the psychological sense and propositions in the logical sense. Roughly speaking, whereas the former understanding considers propositions as mostly investigated by empirical psychology, the latter—is intended as referring to them as belonging to the reality comprising items being the object of logic.

In general, intensional entities, if conceived as existing outside of space and time, serve as various specifications of the concept of propositions in the logical sense; historically, these ideas were initiated by Bolzano's *Sätze an Sich*, Meinong's *Objektive*, Husserl's noemata, Frege's thoughts or Carnap's intensions. Even if we disregard the purely ontological conception of propositions as classes of possible worlds in which sentences are true, the term 'proposition' refers to a dauntingly broad and not quite homogenous variety of entities. This fact is at once noted by the critics of propositions, who point out that the category is too unclear to serve as an effective analytical device in philosophy. Another typical criticism—made by Quine, for example—rejects the excessive ontology of propositions. Although personal philosophical taste generated this objection, it may be quite interesting to observe that even philosophers positively inclined towards empiricism are ready to acknowledge the abstract nature of propositions in order to save them for being used in semantics and other philosophical subjects. According to Carnap (Carnap 1947, p. 25), propositions are objective, non-mental and extra-linguistic entities. Another argument for propositions as separate *sui generis* objects proceeds as follows. Assume that propositions explain how and why sentences are synonymous. Now we should ask for sources of the identity of propositions. A closer inspection shows that criteria of propositional identity are either unclear or essentially rely on the fact that related sentences are synonymous.

Sentences form a grammatical category. It is assumed in grammar that sentences are correct (well-formed in the logical terminology) if they fulfil definite syntactic criteria, for example, they have a subject-predicate form or consist of words arranged in a specific order. The grammatical syntax also determines the mood of a sentence as being declarative, a command, or a question. Grammarians do not need to consider syntactic criteria as either necessary or as sufficient conditions of sentential correctness. Ultimately, sentences are uttered in order to communicate something to someone; the correctness of the utterance is related to its ability to fulfil this undeniably pragmatic task. Successful communication is independent of purely grammatical correctness, at least to some extent. For instance, declaratives can also express, depending on a definite context, questions or commands. For example, the sentence ‘Here is dangerous dog’ expresses the command (or warning) ‘Do not enter!’. We engage rhetorical questions to function as assertions, for example, ‘Do you deny that Heidegger is an obscure philosopher?’. A sentence can be clearly ungrammatical, but yet completely understandable. Although there are notations or jargons, like Morse’s alphabet or logical symbolism, which do not tolerate ambiguities, vagueness, uncertainties, and other unstable properties of natural language, a general syntactic definition of a sentence applicable to all linguistic situations seems impossible. On the other hand, truth and falsehood enters grammar very soon, because declaratives are defined as true or false sentences. In this definition, declaratives are considered as sentences in the logical sense. Truth and falsehood are not attributes of questions and commands. In fact, the role of truth and falsehood in defining declaratives is an argument for taking sentences as truth-bearers. However, this evaluation is only preliminary.

Now, even if we restrict our task to special notations or portions of natural language and assume that we have a syntactic criterion for the predicate ‘being a sentence’, the next problem soon arises. Sentences can be interpreted as either token or as types. On the former interpretation, the sentences ‘snow is white’ and ‘snow is white’ function as two numerically different truth-bearers, even though they seem to be two instantiations of the same entity, that is, a type. Linguistic types in this sense might be defined as abstract categories based on the relation of equiformity of expressions. Unfortunately, this proposal cannot be taken strictly, for possible exceptions in special cases. Firstly, it does not apply, at least not directly, to (a) ‘snow is white’ and (b) ‘Snow is white’, although they seem to fall under the same pattern. That semblance is confirmed by the fact that if we ask a computer to find all occurrences of the sentence ‘snow is white’, for example in this book, it will display instances of (a) and (b). That means that empirical criteria of grammatical syntactic similarity are easily accessible and applicable.

The last statement might suggest that the difference between sentences-tokens and sentences-type is not as important as is usually claimed. Thus, we can say that the sentences (a) and (b) are counted as belonging to the same syntactic category not because of equiformity and typicality but due to their conformity to the same prototype. In the case of languages defined very strictly, prototypes can be reduced to types understood as abstraction classes. The idea of syntactic prototypes affords a real opportunity to defend the conception of truth-bearers as sentences without reference to abstract entities. However, independently whether we use sentences as tokens or types, another issue seems much more relevant. Are there any captivating reasons to prefer propositions as truth-bearers? If the term 'proposition' refers to a linguistic entity, there is practically no difference between sentences and propositions as truth-bearers. Propositions are then only sentences of a certain sort, for example, declaratives or even perfect declaratives—lacking the various linguistic instabilities, mentioned earlier (see Twardowski 1900 for such a conception of propositions; see Chap. 3, Sect. 3.7(C)). Thus, the real theoretical problem arises, when propositions are understood as entities radically different from sentences, but still playing the role of an explanatory category to elucidate how and why sentences are meaningful. I find the following to be the only sound reason for defending the propositional theory of truth-bearers. This reason assumes that sentences are purely syntactic or even solely physical entities (traces of chalk on blackboards, acoustic waves, traces of ink on paper, etc.).

Is the above understanding of sentences the only possible solution? I would answer: no. I suspect that most philosophers, in particular, in the Anglo–Saxon world, believe in a purely syntactic and physicalist theory of sentences. I cannot judge whether the word 'sentence' has this association in English. Anyway, the relevant entry in Hornby's *Advanced Learner's Dictionary of Current English* suggests that it does. We read that

> sentence [is] the largest grammatical unit, consists of phrases and or/clauses used to express a statement, question, command, etc.

Although this explanation is not without ambiguities, it seems to assume that meanings are external to sentences. The Polish term *zdanie* calls for other intuitions that intimate nothing against uttering meaningful sentences outside of purely syntactic criteria. The standard Polish view distinguishes the material side and the semantic or semiotic side of every expression. It allows us to say that basically every expression is endowed with meaning. If this view is adopted, various difficulties of the sentential theory of truth-bearers disappear, or can be neutralized to some extent. Since other theories of meaning also are subject to objections, I take the philosophical liberty to accepting the view that sentences are truth-bearers.

One point should be particularly stressed. I do not maintain that the concept of linguistic expression, as described above, solves the philosophical problem of what meaning is? This is not its task. The distinction between the two sides or aspects of expressions confers legitimacy on a certain theory of truth-bearers and does not serve as a universal philosophical key. I repeat once more that the fundamental objection against taking propositions (or other intensional entities) to be truth-bearers is that they are identified exclusively by the sentences through which they are expressed. The sentential theory of truth-bearers offers the simplest resolution of the issue discussed above. A good outcome of this theory is that sentences are seen as tokens, and prototypes are simply perceived as objects, and these prototypes as such possess the required syntactic category directly. In contrast, propositions, which are not accessible to direct perception, inherit their (syntactical) categorical status only derivatively from corresponding sentences.

(**DG5**) My own inclinations strongly favour the sentential account of truth-bearers. Independently of my personal opinion, this account of truth-bearers also agrees with Tarski's own standpoint. In fact, he alternated between sentence-token and sentence-types. At first, under Leśniewski's influence, he opted for the former alternative (see Tarski 1956, p. 62), but later he definitely preferred to take sentence-types as the basic category. Tarski was fully aware of various defects in the sentential theory of truth-bearers. In particular, he pointed out that various difficulties arise when infinite classes of sentences are admitted (Tarski 1933, p. 174, note 2):

> For example, the following truly subtle points are here raised. Normally expressions are regarded as the products of human activity (or as classes of such products). From this standpoint the supposition that there are infinitely many expressions appears to be obviously nonsensical. But another possible interpretation of the term 'expression' presents itself: we could consider all physical bodies of a particular form and size as expressions. The kernel of the problem is then transferred to the domain of physics. The assertion of the infinity of the number of expressions is then no longer senseless although it may not conform to modern physical and cosmological theories.

Tarski, in his letter to Popper of January, 2, 1955 (The Hoover Institute, box 27, folder 27), clearly preferred 'sentence' over 'statement' as far as truth-bearers are concerned. Popper (see Popper 1955, p. 333, note 1) replied:

> I understand that Tarski prefers to translate '*Aussage*' and '*Aussagefunktion*' [Tarski's remark concerned translation from German – J.W] by 'sentence' and 'sentential function' (while I am using here 'statement' and 'statement function' […]).

These quotations settle the problem of how Tarski himself understood truth-bearers in **STT**.►

(**DG6**) When Tarski spoke about expressions as products of human activity, he probably alluded to Twardowski's distinction between actions and products (see Twardowski 1912), which became very influential in Poland. Twardowski's ideas open some possibilities for the solution of difficult problems concerning truth-bearers. According to Twardowski, language is a product of certain mental acts. In contemporary terminology, this means that linguistic expressions just supervene on human acts of a particular kind. These acts can be interpreted as conferring sense (meaning) on objects selected as bearers of semantic properties; the acts in question are similar to signifying operations in Husserl's sense. The linguistic expressions are not only supervenient, but are also durable—that is, their existence persists beyond the existence of the corresponding creative acts. This property of expressions provides a way to a solution of how set of sentences can be infinite when the initial collection of linguistic items is finite. It defines an infinite (infinity is understood here as potential) set of sentential expressions as being the smallest set containing the linguistic products produced a specific time (this set is finite) which is closed under standard logical operations consisting in the application of logical connectives and quantifiers (see Chap. 5, Sect. 5.2A). The propositional account of truth-bearers has no problem with infinity, because the number of propositions is infinite by definition.►

(**DG7**) Contrary to the opinion of many philosophers, the sentential conception does not solve the problem of meaning in a satisfactory way, at least if language is understood as a product of mental or psychophysical human activities. Since we can create only a finite number of linguistic expressions, it is actually puzzling how the infinite number of propositions can acquire meaning (see definition above). Let me also remark that all problems concerning the distinction of sentence-tokens and sentence-types automatically apply to thoughts, statements, beliefs, etc. as truth-bearers, provided that they are understood as actual psychical events. If these entities are understood as being concrete, they are numerically different from others items of a related kind. For example, assume that my thought m constitutes the meaning of an expression e and the thought m' constitutes the meaning of the expression e', where e and e' are recognizable as exemplars of the same prototype. Since m and m' are different, I need to say that they are instances of the same type of thought. However, the concept of the prototype of thought has no clear sense. In particular, we have no precise idea how to count thoughts as falling under thought-prototypes. Thus, the concept of prototype of thought is plagued by the same difficulties as all abstract objects.►

(**DG8**) There still remains one problem concerning truth-bearers, namely the question of their scope. As I already noted, elementary grammar divides sentences into declaratives, questions and commands. By definition, questions and commands are excluded from the domain of truth-bearers, that is, sentences in the logical sense, because they are neither true nor false. A simple test whether A is a truth-bearer runs as follows. Take the expression A and precede it by the phrase 'it is true that' (or 'it is false that'). Check then, whether the whole expression 'it true that A' is correct or not. For example, correct are (a) 'London is a city in England' and (b) 'Paris is a city in Germany', but (c) 'Is Berlin the capital of Germany?', (d) 'Close the door' and (e) 'London' are obviously incorrect. In particular, example (e) shows that our test also works for expressions other e than sentences. Roughly speaking, it restricts the set of truth-bearers to the set o declaratives, independently of whether they are true or false. However, we also have a controversial category of sentences, namely value-sentences and normative sentences (norms). The outlined test qualifies them as truth-bearers, at least from the intuitive point of view. Phrases, like (f) 'it is true that a is good (beautiful, etc.)' or (g) 'it is true that everybody should pay taxes' are obviously correct. Yet many philosophers deny that norms and evaluations are true or false. Clearly, grammatical argumentation does not help in this case. One must argue some other way to justify that value-sentences and norms are neither true nor false. In fact, arguments for excluding value-statements and norms from the domain of the true-or-false appeal to general philosophical views, on the nature of such sentences, emotivism, for instance.►

4.3 Is Truth an Epistemological or Ontological Concept?

Although truth is one of the most frequently investigated concepts in epistemology, it is also frequently employed in ontology (see Chaps. 1, 2 and 3 for historical examples). As far as epistemology and ontology are concerned, my treatment of truth-bearers definitely favours the epistemological concept of truth. It was Aristotle (see Chap. 1, Sect. 1.4), who pointed out that truth is in the mind, not in things. Yet the ontological concept of truth cannot be neglected (see also Woleński 2004d, Woleński 2013). Roughly speaking, under ontological understanding of truth, truth is an attribute of being. This account of truth is legitimised to some extent by the claim that:

(19) It is true that *A*, can be replaced by
(20) It is a fact that *A*.

Now, if we accept (a) that (19) and (20) are equivalent, and (b) that being is the totality of facts, we obtain something very similar to the Scholastic principle *ens et verum convertuntur* (being and truth are convertible; see Chap. 2, Sect. 2.3). Clearly, being cannot be false, because false being is not being at all. Hence, true being is simply being. This simple observation shows that the adjective 'false' in the phrase 'false being' is a modifier (modifying adjective), that is, one which alters the meaning of the noun to which it attached; well-known examples of phrases with adjectives 'false' and 'dead' as modifiers are these: 'false gold', 'false friend' or 'dead men'. Consequently, the adjective 'true' in the phrase 'true being' may be considered redundant, because it does not add anything to the word 'being' in the sense that it not to refer to any property of being. 'True' in this role is a redundancy-predicate operator, that is, the expressions 'true P' and 'P' are equivalent. If we work with ontological truth-bearers, it is clear that being cannot be a falsehood-bearer. This violates the principle (*) from (**DG3**). On the other hand, the status of sentences as truth-bearers is independent of whether they are true or false. This confirms an earlier remark that the term 'truth-bearer' is an abbreviation for 'truth-or-falsehood-bearer' or even 'logical-value-bearer'. More importantly, this analysis gives a fairly strong argument for the view that the epistemological and ontological concepts of truth are essentially different.

(**DG9**) In speaking about 'true' in 'true being' as a redundancy-predicate operator, I do not want to suggest that its role is trivial on other occasions. For example, if someone says 'it is a true work of art' or 'it is true gold', he or she might try to convince someone else about the value of this or that piece of art or the authenticity of a gold coin. However, such uses have rather rhetoric import than epistemological significance. Let me also note that the phrase 'true being' can indicate that someome had in his/her mind 'the most true' being. Perhaps Platonic forms are true being(s) in this sense. Similarly, 'false being' cane refer to something less real or worse as compared with something else.►

(**DG10**) One could defend the equivalence of the epistemological and ontological concepts of truth in the following way. Let us replace the principle *ens et verum convertuntur* by *factum et verum convertuntur*. The sentence (a) 'London is in China' is false. This may be expressed by (b) 'It is not a fact that London is in China'. However, sentences like (b) do not express non-being, unless we form strange utterances, like 'being London in China is not being' or 'that London is in China is not being'. The simplest way to express that it is not a fact that London is in China consists is to say 'London is not in China'. The sentence 'London is in China', although false, concerns the real world in an equally literal way as the sentence 'London is in England'. See Kastil 1947 and Hofstadter 1965 for further comments about the ontological concept of truth.►

4.4 The Problem of Truth-Definability

Some philosophers maintain that truth-definition is dispensable at all and should not be defined (compare the impressive title of Davidson 1996—"The Folly to Trying to Define Truth"). This is the view of theorists who defend redundantism, minimalism, deflationism or disquotationalism or a view that 'is true' is a primitive idea (see Chap. 3, Sect. 3.8 and Chap. 9, Sect. 9.7). These theories reduce everything what is important for the truth-theory to the formula (**T**-scheme; I quote once again).

(21) A is true if and only if A.

Another position is adopted by so-called substantive theories of truth (see Baldwin 1991, Sher 1999). It claims that truth should be defined in a material or real manner, that is, by pointing out a property (correspondence, coherence, evidence, consensus etc.) that is not intelligible or badly understood without a definition (see also the classifications of truth-definitions mentioned in this section below).

(**DG11**) (A historical digression supplementing Chap. 3, Sect. 3.2) Early Moore (see Moore 1899) considered truth as a relation that is not definable but should be characterized as recognizable. Cartwright (see Cartwright 1987a, p. 71) suggests that this view is parallel to Moore's famous analysis of goodness as a simple, non-analysable quality. However, Moore himself never invokes such a comparison. Cartwright ascribes the same view to Russell, in particular, by reference to a fragment from Russell 1903. However, this interpretation seems incorrect. In fact, Russell (see Russell 1903, p. 3) said that mathematics considers the concept of truth (together with implication and membership) among primitive ideas. Russell (see pp. 35, 38) mentions problems that show up in the definition of truth, but regards them as global, not local, that is, as related to mathematics itself. These remarks give no basis for the contention that Russell considered truth as indefinable. In fact, he offered several definitions of this notion in his writings (see Chap. 3, Sect. 3.2).►

The considerations in Sects. 4.1–4.4 suggest that a possible substantive definition of truth should assume something about truth-bearers, follow either the epistemological or ontological path as well as remain connected to problems I–X, or at least to some of them. Sher (see Sher 1999, p. 147) remarks that the prevailing substantive definitions of truth identify a single universal principle or a collection of such principles in order to capture a general and intuitive factor which is particularly important for the truth-concept. Although she does not rule out that this strategy might be successful, she point out that (a) such an universal factor is not necessary for a substantive theory; (b) various trivial principles (for example, the scheme (21)) cannot serve as the basis; (c) the multidimensionality of the truth-problematic makes the chance of finding such a factor rather small; and (d) the existence of the factor in question cannot be posited in advance, but has to be established or rejected as a result of essential philosophical analysis. Thesis (a) may be confirmed by Frege's conception of truth. Although it dispenses with a truth-definition, we cannot declare that it is not substantive. Further, minimalism and other views that focus on (21) do not help much in solving most of the problems in truth-theory. This observation provides some evidence for (b). The collection of problems (I)–(X) actually attests to how multidimensional and complex the problem of truth is. If a truth-definition were to be used as the universal key for everything in truth-theory, scepticism concerning the chances of finding a defining formula would be undoubtedly well justified. Yet if the universality-claim (in Sher's sense) is rejected, the matter does not look so hopeless. Thus, the multidimensionality of the concept of truth is not at odds with a possible success of defining truth. Finally, although (d) seems reasonable, many philosophers forget this claim, especially when initiating polemics with the view of others (see Chap. 9 for documentation of this statement).

As far as how to classify truth-definitions is concerned, I will report three proposals. At first, I outline the framework given in Ajdukiewicz 1949 (since Tarski was educated in the spirit of this classification, it somehow contributes to a better understanding of **STT**). Ajdukiewicz divided all truth-definitions into two groups. The first group has only one element, namely the classical truth-definition. According to the popular version of this conception, truth consists in agreement between thought and reality (see Chap. 3, Sect. 3.7(G) for an improved formulation; I still use the labels 'the correspondence theory of truth' and 'the classical theory of truth' as equivalent, but see Chap. 3, Sect. 3.7 on the rejection this identification in Poland). The second group includes so-called non-classical truth-definitions. They fall under the scheme (p. 12):

(**NC**) Truth consists in agreement of thought with final and irrevocable criteria.

Ajdukiewicz lists four cases that fall under (**NC**) depending on what is referred by a criterion **C** considered as final and irrevocable (there is, of course, problematic what these predications refer to; see the next section for some remarks on this issue). The first criterion is coherence. It proposes to define truth by the internal coherence of thoughts. Secondly, we have the common agreement theory:

statements (the term ‘statement’ occurs in Ajdukiewicz) are true if and only if they are subjected to common consensus. The third criterion invokes evidence: thoughts are true if and only if they are evident; the Neo-Kantian view that truth is determined by transcendental cognitive norms is a special case of applying the evidence criterion. Finally, the pragmatic criterion establishes truth as a derivative of utility. Every criterion generates a separate truth-theory. Accordingly, we have four non-classical truth-conceptions: coherence theory, common agreement theory (the consensual theory), evidence theory, and pragmatic theory (the utilitarian theory); the labels in parentheses do not occur in Ajdukiewicz. His list gives some general ideas of how truth can be defined, how to compare particular definitions or establish various relationships between them (for example, that they are not all mutually contradictory; in fact, a thought can be evident and commonly accepted), and how to look for clues for arguing in favour of particular proposals.

Second, Richard L. Kirkham (see Kirkham 1992, pp. 20–31) distinguishes (I limit the description only to the most general level) the metaphysical project (for instance, the correspondence theory), the justification project (for instance, coherentism) and the speech-act project (‘true’ as a performative word). Third, Künne (see Künne 1985) divides truth-theories into epistemic (truth is recognized by accepting something as true) and non–epistemic (truth is not reducible to epistemic circumstances); the latter conception can be either relational or non-relational). Although some concrete truth-theories can be variously classified, others appear as fairly paradigmatic. For instance, pragmatic and consensualist theories are epistemic, the minimalist is non-relational, but, on the other hand, the classical and coherence definitions (under special constraints) are relativist (this classification is modified and somehow extended in Künne 2003, Chap. 1, but I will not discuss the new version). Künne’s proposal can be very easily compared with the substantivism/minimalism distinction. We can identify minimalism with non-epistemic, non-relational approaches and the remaining rubrics in Künne’s classifications with substantivism. All places occurring in Ajdukiewicz’s map regard truth as a substantive property, the classical definition is relational and non-epistemic, and non-classical theories are epistemic, but their relational character depends on something else. If we say that they define truth via its relation to some criteria, all are relational. However, taking into account Russell’s view on being relational, the coherence approach is not such.

4.5 Truth-Definitions and Truth-Criteria

Ajdukiewicz’s division of theories of truth into classical and non-classical ones immediately leads to the problem of how truth-definitions and truth-criteria are mutually related. In fact, the scheme (**NC**) from the last section generates various criteria-determined truth-definitions, that is, identifications of the nature of truth with the various possible ways of demonstrating whether a given truth-bearer is true or not (see Huby, Neal 1989 on truth-criteria in antiquity). Presumably one could claim in advance that a sharp distinction between truth-definitions and truth-criteria should be introduced as an additional adequacy constraint in building a satisfactory account of truth. However, this claim would break Sher's condition (d), which I consider to be fully reasonable. Although many concepts are defined independently of possible criteria of their identification (for example, the concept of the discoverer of Africa illustrates the issue: we know what it means to be the discoverer of Africa, but we have no idea how to identify this person), nothing a priori decides that the notion of truth belongs to that variety. The matter must be investigated more closely, without any a piori presumptions.

According to tradition, the classical truth-definition is not bound by criteria. This opinion seems correct and I accept it. There are at least two reasons to favour truth-definitions based on criteria over the classical account. The first tries to employ arguments going back to the ancient sceptics who argued against any truth-criterion. In particular, as the sceptics said, such a criterion can be either direct or indirect. The first consists in perceiving a thing, but the second relies on inferences. The direct criterion does not work, because the senses are not reliable (I skip more detailed arguments, for instance, Agrippa's tropes). In particular, for any assertion A based on the senses, A and not-A are isotonic (isostenic), that is, both have equal or at least approximately similar perceptual evidence. Hence, we have no way to decide which statement of the pair {A, not-A} is true. On the other hand, the indirect criterion is burdened by the defects of *petitio principii* or *regressus* ad infinitum. Now, let the truth of A be checked by an inferential procedure. Since the case of deduction is particularly instructive here, we must find A' such that (a) A is deductively inferable from A', and (b) A' is true. Condition (a) defines the indirect criterion, but condition (b) must meet the *petitio principii* objection because we have the question of the grounds on which the truth of A' was asserted. In order to avoid *petitio principii*, we have to appeal to a true A'', from which A' is inferable from. Clearly, this step is repeatable with respect to the assertion A''. According to the sceptics, *regressus* ad infinitum or *petitio principii* cannot be avoided in the case of any indirect criterion of truth, whether deductive or inductive, Descartes believed that his evidence account of truth successfully met the sceptics' challenge, although it is unclear whether he rejected classical theory (see Chap. 2, Sect. 2.5). Brentano (see Chap. 3, Sect. 3.5) wanted to combine evidence theory with a suitably reinterpreted classical truth-definition. This is not the place to discuss whether these attempts were successful. I only wish to call attention to the general strategies of arguing for non-classical solutions in aletheiology.

Other criterion-based theories of truth usually ignore scepticism. It is now time to make some remarks about Ajdukiewicz's classification of truth-theories. His formulation of (**NC**) is too radical. He no doubt exaggerated when he said that truth-criteria appeal to ultimate and irrevocable standards. It is false with respect to most versions—contemporary ones, in particular, of coherentism, consensualism and utilitarianism. The defenders of coherence, like Bradley or Neurath 1931 (see also Hempel 1935), of consensus, like Jürgen Habermas or Thomas Kuhn, or of pragmatism, like James or Dewey, did not propose ultimate and irrevocable criteria. On the contrary, most of them accepted fallibilism and relativism. The main idea justifying non-classical definitions was (and still is) that the average person is much more interested in ways to check truth than in the concept of truth itself. The possibility to distinguish truth from error, lies or falsehood is of the utmost importance in science, daily life and legal matters. This attests, according to many philosophers, to the demand that criteria-determined truth-definitions better address real intuitions and cognitive needs than the classical theory.

One could say that sceptical arguments also go counter non-classical truth-definitions, but this contention is not careful sufficiently. It is often neglected that the ancient sceptics challenged the concept of knowledge as *episteme* in the sense of Plato and Aristotle. The sceptical arguments opposed the truth-criterion for *episteme*, assuming that truth consists in agreement of cognition with reality. However, the contemporary criterion-advocates do not define knowledge as *episteme*. Consequently, it seems sceptical arguments do not apply to definitions of truth as working within a framework of accounts of knowledge other than *episteme* in its classical setting. Prima facie, it seems that if the relativity of knowledge is admitted, most sceptical arguments are not applicable, but if perception serves as the ultimate and certain criterion of knowledge, scepticism must be taken very seriously, independently of whether the classical theory is assumed or not. It would appear that just *episteme*, not truth, constitutes the main problem. If the correspondence theory of truth is abandoned, truth-criteria lose their function as a test of the agreement between truth-bearers and reality, and should be evaluated according to other standards, for example, that offered by the coherence theory of justification.

Finally, I would like to note a specific problem pertaining to truth-criteria that plays an important role in my further considerations. I already said definitions of many concepts do not appeal to criteria. I can now enter more deeply into the problem in question by using an example from metamathematics. Consider a proof of a sentence A via the ω-rule. This rule admits an infinite number of premises. The problem is not trivial, because proof of the sentence 'The Peano arithmetic of natural numbers is semantically complete' requires the ω-rule. We can give a precise definition of what it means that something is provable via the ω-rule, but there is no criterion to check step by step whether such a proof is correct, because we cannot effectively operate with infinite sets of premises. The qualification 'effectively' is absolutely crucial in this respect, not only in logic or mathematics, but whenever the concept of criteria is invoked. In particular, it seems that in formulating of the non-classical truth-theories, philosophers tacitly make two important assumptions, namely, (a) that the truth-criteria in question are effective, and (b) that the classical truth-definition is at odds with effective truth-criteria. It is very difficult to say precisely how the effectiveness of truth-criteria is conceived in particular non-classical truth-theories, but the problem should not be disregarded. I will investigate this question with respect to coherentism in Chap. 9, Sect. 9.8. Anticipating one of the main philosophical outcomes of this book, I believe that an important interpretative consequence of **STT** as a formal truth-theory consists in demonstrating that the classical concept of truth is of an infinite nature and therefore cannot be exhausted by any finitary truth-criterion.

4.6 How Many Truth-Values?

The definition of sentences in the logical sense as true or false immediately suggests that there are two and exactly two logical values. If the principle of bivalence, that is, the statement

(**BI**) Every sentence in the logical sense is either true or false,
functions as a general metalogical principle, then the set of truth-bearers (**TB**) is exhaustively and strongly disjunctively (either, or) divided into the set of truths (**TR**) and the set of falsehoods (**FL**) (for simplicity, I will omit 'strongly' in the subsequent considerations). More formally, the following principles hold:

(22) (a) $\mathbf{TB} = \mathbf{TR} \cup \mathbf{FL}$;
(b) $\mathbf{TR} \cap \mathbf{FL} = \varnothing$.

Formula (22a) asserts that the division of truth-bearers into truths and falsehoods is exhaustive; this statement expresses the metalogical principle of excluded middle (**MEM**), but (22b) says that no truth-bearer is simultaneously true and false (truth and falsehood are exclusively disjunctive); this statement expresses the metalogical principle of contradiction (**MC**). Jointly, (22a) and (22b) yield the principle of bivalence (**BI**), which is the conjunction of (**MEM**) and (**MC**); this account was proposed by Łukasiewicz. Since (**BI**) is a characteristic metalogical principle of classical logic, this logic determines the logical behaviour of truth-bearers under the assumption that their division into truths and falsehoods is exhaustive and disjunctive (see Wansing, Schramko 2011 for a very abstract approach).

Two logical values are the lower bound of the number of logical values for any consistent logic. Assume that there is only one logical value $\mathbf{v}$. Accordingly, we have that for every A, $v(A) = \mathbf{v}$ (read: the logical value of A is equal to $\mathbf{v}$). In particular, (a) $v(A) = v(\neg A)$. Then, $\mathbf{v}$ must be also the designated value, because there is no other possibility. Now, since a tautology is a formula that has the designated value for any possible evaluation of its subformulas (this is an informal and simplified explanation; see Chap. 5, Sect. 5.2 for a more detailed treatment), every sentence is tautological, and the simplest logical law (b) $A \Rightarrow A$ enables us to prove an arbitrary sentence. Imagine (a) and its predecessor, that is A. Since A has the designated value, the detachment rule allows us to assert A. Hence, every A is a theorem and our logic is inconsistent in the absolute sense; a set of sentences is inconsistent if and only if every formula belongs to its consequences. The same follows from (a) a pair of contradictory sentences having the same designated value. Thus, a single-valued logic does not obey the fundamental metalogical requirement. Perhaps Hegel would accept such a construction under his view that Being is identical with Non-Being, but even this is uncertain.

(**DG12**) Bradley, who was rather close to Hegel, considered every sentence as a mixture of truth and falsehood, but this does not entail that there is only one value. In fact, Bradley advocated the idea of degrees of truth. Although it is unclear how this idea could be formalized (perhaps as a probabilistic logic), it certainly does not entail that there is only one logical value. The view known as dialetheism admits true contradictions, but its logic has two values at least. Although dialetheism rejects the logical principle of contradiction, this philosophy of logic is not opposed to the metalogical one. The logic of dialetheism is paraconsistent, that is, it admits contradictions but prevents absolute inconsistency by omitting the formula $(A \wedge \neg A) \rightarrow B$ from its stock of logical principles. Because of this, paraconsistent logic avoids overfullness (sometimes called 'explosion'), that is, taking all formulas as theorems. This logic is not a single-valued (see Béziau, Carnielli, Gabbay 2007 for an encyclopaedic survey of paraconsistency).►

Dividing truth-bearers into sub-classes is one problem, but how far this can be carried out is a far more complex issue. It is strongly connected with many-valued logic and related constructions. Many-valued logic is based on rejecting (**BI**), more precisely on dispensing with (**MEM**). Assume that we have n logical values $\mathbf{1}$, $\mathbf{2}$, …, $\mathbf{n}$; I, for the sake of simplicity, only consider the finite case only (moreover, the symbolism in this section is only temporary). There is no difficulty with (**MC**). Suffices to say that there is no A such that $\mathbf{v}(A) = \mathbf{i} = \mathbf{j}$, where (a) $\mathbf{i} \neq \mathbf{j}$ and (b) $\mathbf{1} < \mathbf{i}, \mathbf{j} \leq \mathbf{n}$. This means that no sentence has two different logical values, but assuming that $\mathbf{1} = \mathbf{t}$ and $\mathbf{n} = \mathbf{f}$ ($\mathbf{t}$ = truth; $\mathbf{f}$ = falsehood), sentences can possess logical values other than truth and falsehood. Generalizing (22a) to:

(23) $\mathbf{TB} = \mathbf{VA_1} \cup \mathbf{VA_2} \ldots \cup \mathbf{VA_n}$,

Preserves the exhaustive division of truth-bearers with respect to logical values $\mathbf{1}$, …, $\mathbf{n}$, where $\mathbf{VA_i}$ is a set of truth-bearers having the value $\mathbf{i}$ ($\mathbf{1} \leq \mathbf{i} \leq \mathbf{n}$). In order to define the concept of truth-bearer in a similar manner as in the case of sentences in the logical sense, we need to say that a truth-bearer (in fact, it is an abbreviation for 'logical value-bearer') is any sentence A which satisfies the condition $v(A) = \mathbf{i}$ ($\mathbf{1} \leq \mathbf{i} \leq \mathbf{n}$). In general, a truth-bearer (a sentence in the logical sense) is a sentence that possesses one of the adopted logical values.

Formula (23) is not equivalent to (**MEM**), unless we adopt a generalized form

(24) For any A, A is true or A is not true,

provided that truth is among the set $\{\mathbf{1}, \ldots, \mathbf{n}\}$, but (24) is not a version of the original (**MEM**). An example based on Łukasiewicz's three-valued logic explains what is going on. Assume that we have three logical values, namely $\mathbf{1}$ (truth), $\mathbf{2}$ (neutral value; neutrum) and $\mathbf{3}$ (falsehood). The valuation for negation is given by (a) $v(A) = \mathbf{2}$ if and only if $v(\neg A) = \mathbf{2}$. Of course, this logic rejects the principle that A is true or $\neg A$ is true, because these sentences can be valued by neutrum. Moreover, since $v(A \wedge B) = \mathbf{2}$, for $v(A) = v(B) = \mathbf{2}$, the logical principle of contradiction fails, because $v(\neg(A \wedge \neg A) = \mathbf{2}$, if both A and $\neg A$ are valued only by neutrum. Yet (**MC**) holds without any restriction. Let me also note that the presence of truth in (24), that is, the generalized principle of excluded middle is entirely accidental, because we can say

(25) For any A, $v(A) = \mathbf{i}$ or $v(A) \neq \mathbf{i}$, for any $\mathbf{1} \leq \mathbf{i} \leq \mathbf{n}$.

This suggests that only the two-valued metalogical principle of excluded middle is not trivial, because falsehood commutes with negation and being non-true, according to:

(26) For any A, A is not true if and only if A is false if and only if $\neg A$ is true.

This shows that one should be very careful with understanding the principles (**MEM**) and (**MC**). In particular, there is a fundamental difference between these principles and the corresponding logical laws of excluded middle and contradiction: $A \vee \neg A$ or $\neg(A \wedge \neg A)$ respectively. If we interpret them by **TR**, analogically as in the formula (22), we obtain

(27) (a) $\mathbf{TR} \cup -\mathbf{TR} = \mathbf{TB}$; and
(b) $\mathbf{TR} \cap -\mathbf{TR} = \varnothing$.

The direct logical reading gives special instances of both tautologies, namely $\mathbf{TR} \vee \neg\mathbf{TR}$ and $\neg(\mathbf{TR} \wedge \neg\mathbf{TR})$, where the symbol **TR** functions as a propositional constant. The formulas (27a) and (27b), eventually their mentioned special instances have a different content than (22a) and (22b), unless we agree in advance that we have only two-values, namely **TR** and **FL** (= **not-TR**) (see also below).

(**DG13**) The three-valued system of Łukasiewicz was the first mature many-valued logical system. In the 1920s Łukasiewicz generalized many-valued logic, firstly to logic with an arbitrary finite number of logical values, and, then to the denumerable infinite logic. There is no problem with constructing many-valued logic with sets of values of arbitrary infinite cardinality. Another direction of research, initiated by Hans Reichenbach, consists in linking many-valued logic with probability. See Zinoview 1963, Rescher 1969, Woleński 1989 and Malinowski 1994 for more detailed historical information.►

Admitting more than two logical values is not the only way to break bivalence. That can also be achieved by introducing truth-value gaps. Although many-valued logics with truth-value gaps are possible, I will mention only the two-value case. The idea is to qualify some sentences as neither true nor false, but without ascribing other truth-values. Thus, a sentence A is said to be a truth-value gap if and only if it is neither true nor false. I will not enter into the technical details of such constructions, which are connected with the semantic method of supervaluation (see Van Fraassen 1971). However, one point must be made. If truth-value gaps are introduced, the parallelism between truth-bearers and sentences in the logical sense disappears. We can still keep the definition of truth-bearers as true or false, but there is no reason to exclude sentences with truth-value gaps from sentences in the logical sense insofar as they are premises and conclusions of correct inferences. Thus, semantics with truth-value gaps must be based on purely syntactical criteria of being a sentence in the logical sense or based on new semantic measures.

The question that now arises concerns reasons to introducing many values or truth-value gaps. Such logical constructions are motivated by some puzzling problems (for instance, see the surveys in Haack 1978, Priest 2008) like (I skip bibliographical references) sentences about the future (Łukasiewicz), non-verifiable sentences about the past (Dummett), undecidable sentences in mathematics (Bochvar), sentences that describe measuring in quantum mechanics (some quantum logicians), sentences leading to semantic paradoxes (some advocates of the so-called naive semantic conception of truth), for example, the Liar sentence 'I am lying', or the sentences with empty descriptions such as 'the present King of France is bald' (Strawson). Consider the sentence 'I will visit my friends tomorrow'. It is a contingent sentence about the future (a future contingent, as it is customarily called). Its contingency consists in a possibility that I will visit friends, but also in a possibility of changing my plans or something happening that prevents me from going to the place where my friends live. Is this considered sentence true or false at present, that is, today? The advocates of many-valued logic or true-value gaps deny that this sentence is true or false. They say, depending on the accepted solution, that it possesses another logical value, for instance, neutrum, or that it is a truth-value gap that is, neither true nor false, without having any other logical value. It is one of the most celebrated philosophical problems. I will return to this matter in Chap. 9, Sect. 9.5.

(**DG14**) There is no agreement among philosophers what consequences for the general theory of truth follow from introducing many (>2) logical values or truth-value gaps. Łukasiewicz, who championed many-valued logic, saw no reasons for abandoning the classical truth-theory. On the other hand, anti-realists inspired by intuitionistic logic propose an approach to truth via assertibility (it will be analysed in Chap. 9, Sect. 9.9). In general, this book favours classical logic, that is, based on (**BI**). Hence, remarks on truth-theories based on other logical devices are only marginal (see Linke 1949 for the view that if truth and falsity are understood ontologically, many-valued logic does not force us to reject them as the only logical values). The adjective 'classical' is used to qualify logic as well as one of the truth-theories. Is it accidental? We will see that, according to **STT**, there is a strong link between the classical theory of truth and classical logic.►

4.7 Is the Division of Truth-Bearers Stable?

The question whether the division of the set of truth-bearers into some distinctive subsets corresponding to logical values is stable, concerns a possibility of changing logical values by truth-bearers. Philosophically speaking, it is equivalent to a more familiar problem, namely, whether sentences possess truth-values in a relative or absolute way. Discussions about the relative or absolute nature of truth have been around at least since Socrates engaged in exchanges with the Sophists (see Wentscher 1941 for a brief historical sketch of older conceptions). I restrict myself here to the case of truth and falsehood as logical values, although the question can also be raised with respect to semantics with more values or truth-value gaps. For example, in Łukasiewicz's three-valued logic, the present neutra (the sentences which have the third logical value at present) appear as truths or falsehoods in the future. Thus, categorising sentences into sentences possessing logical value is not a stable method, because it sometimes happens that a logical-value bearer changes its previous status. On the other hand, Łukasiewicz recognized that truths and falsehoods are stable in the sense that what is true (false) cannot change its logical value (see Chap. 3, Sect. 3.7(C)) and Chap. 9, Sect. 9.5 for further remarks). In other words, the change of logical values is doubly restricted. Firstly, only neutra are changeable. Secondly, the change is directed in the sense that it goes from now to the future.

If we turn to bivalent semantics with truth and falsehood as the only possible logical values our problem is reduced to the question of whether truths can become falsehoods or vice versa. Absolutism in truth-theory denies that it might can, but relativism admits that it is possible. Relativism points out various circumstances in which sentences can change their logical values. Truth (the same applies falsehood) depends on time, space, culture, utility, theory, justification, the acquired level of knowledge, fuzziness, attitudes, etc. Perhaps the case of time is the most important. Assume that $A \in \mathbf{TR}$. Consistency requires that $A \notin \mathbf{FL}$. How to preserve consistency if we take relativism for granted, that is, if something true may change its initial status and move the domain of falsehoods or, reversibly, something initially false convert into truth? The only possibility seems to consist in indexing logical values using a temporal index. Let the indexes s and t refer to different moments of time. We say that the formula $A \in \mathbf{TR_t}$ means 'A is true at time t' and that the formula $A \in \mathbf{FL_s}$ means 'A is false at time s'. The relativist says it is possible that $A \in \mathbf{TR_t}$, but $A \in \mathbf{FL_s}$. The absolutist rejects this view and says that truth-bearers are stable with respect their status in their given division. This explains why the adverb 'simultaneously' appeared in comments about (**MC**) in the previous paragraph.

The relation of particular truth-theories to the issue of absolutism and relativism is fairly complex. The classical conception is commonly considered as absolutist. The utilitarian theory and the consensus theory are relativistic, because the status of consensus or ascriptions of utilities to sentences (in this case, rather judgements) as guides of action are changeable. The evidence theories of Descartes and Brentano are absolutists; the coherence theory in some versions (for example, that of Bradley) is absolute, but Neurath's or Hempel's coherentism is relativistic. It is important to ask whether a given truth-theory implies absolutism or relativism (see Chaps. 1, 2, 3 and Chap. 9, Sect. 9.5).

4.8 Truth and Logic

The problem of truth was always considered as closely related to logic. Frege was probably the first logicians to discuss this question in a sufficiently general way; recall that he considered logic as explicating the laws of truth. Continuing my earlier remarks about assertion and assertibility (see Chap. 3, Sect. 3.6), I am hesitant to regard assertibility as a logical concept (see Weaver 2015 for an extensive analysis of mathematical assertibility and its relation to truth). The adjective 'logical' can be applied less or more broadly. On the first interpretation, logical concepts are only those studied in logic and metalogic by purely formal devices, provided by logic itself. If 'logical' is understood in a in a broader sense (as it was customary in Frege's times), it refers to various concepts employed in logic in the broader sense, for instance, semantic ones (see Chap. 5, Sects. 5.1 and 5.2 for further remarks on the concept of logic).

Let me explain what is involved by taking the notion of truth as an example. The concept of logical truth can be regarded as logical, because it finds its full explication in metalogic. Firstly, we define the concept of logical theorem. Then we define logical truth as a formula that is true under all interpretations (models). Thus, the concept of logical truth acquires an exact meaning. However, one can say that the same applies to the concept of truth *simpliciter*, in particular, when **STT** is taken into account. An additional argument points out that since the truth under all interpretations assumes the concept of truth under a given interpretation, the concept of logical truth assumes that truth *simpliciter* was already defined. My main argument for considering logical truth as a logical notion relies on the fact that it is a semantic counterpart of the concept of logical theorem. It does not hold for truth *simpliciter*. An analogous discussion can be conducted with respect to Frege's idea of developing the concept of truth by means of logic (see Chap. 3, Sect. 3.6). On the other hand, since the boundary between ranges of 'logical' in the two interpretations is vague, there is room for various standpoints. Perhaps we should employ two adjectives, namely, 'logical' and 'formal', in such a way that the latter refers to studies using devices from logic and mathematics. Thus, 'assertibility', refers rather to a formal than logical concept, unless it is defined as logical provability that is, asserted on purely logical grounds.

Is Frege's view that logic is the science of true the only possibility? Surprisingly enough, the answer is no, because it is not difficult to define the consequence relation as preserving falsehood. I give only a brief account here (see also Woleński 1995). Consider the formulas

(28) (a) $A \wedge B \Rightarrow A$;
(b) $A \Rightarrow A \wedge B$.

The first is truth-preserving. Assume that $A \wedge B$ is true. Thus, A is true and B is true —so A is true. Assume that A is false. Thus, $A \wedge B$ is false too. Summing up, it is impossible for the antecedent of (28a) to be true if its consequent is false. The preceding sentence gives another version of the idea of truth-preservability. The situation is different in the case of (28b). Clearly, A can be true, but $A \wedge B$ false if B is false, so (28b) is not truth-preserving. However, it preserves falsehood, because if A is false, $A \wedge B$ is false as well. Further, $A \wedge B$ cannot be true, provided that A is false, which means that it is impossible for the antecedent of (28b) to be false, but its consequent to be true. Thus the formula (28b) is not truth-preserving, but its falsehood preserving. This observation motivates to considering falsehood as so-called designated value that is preserved by correct logical inferences.

Consider now the formulas:

(29) (a) $A \Rightarrow A \vee B$;

(b) $A \vee B \Rightarrow A$.

A simple analysis shows that (29a) preserves truth, but (29b) preserves falsehood. Now, (28a) and (29b) are mutually dual (the symbol $\wedge$ in (28a) is replaced by the symbol $\vee$ in (29b)); the same goes for (28b) and (29a). Thus, (28b) and (29b) are examples of principles of dual logic, that is, logic preserving falsehood (see also (**DG13V**)). The duality of $\wedge$ and $\vee$ is syntactic. It is parallel to the semantic duality of truth and falsehood. We can say that if the antecedent of (29a) is assertible, the relative consequent is assertible on purely logical grounds. However, this description cannot be applied with respect to (28b) and (29b). Instead we should say that if the antecedents of (28b) and (29b) are rejectable, then their consequences are rejectable. Thus, assertibility and rejectability are the next duals. If assertion is conceived as something factual, then possibly everything happens to be asserted. On the other hand, assertibility is, so to speak, something capable of being asserted in the logical sense.

The assumption that only truths are assertible is perfectly sound, just similarly as its dual assumption that falsehoods are rejectable, is also correct. Yet we have minimal assertibility and minimal rejectability as related to tautologies and contradictions. It is stipulated by

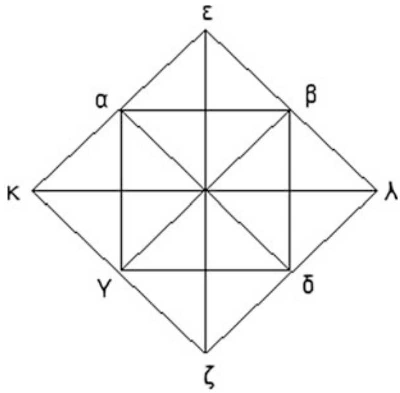
(30) (a) $\mathbf{As}^{\mathbf{m}}$ (A) if A is a tautology;
(b) $\mathbf{Rj}^{\mathbf{m}}$ (A) if A is a logical contradiction (counter-tautology).

The formula (30a) says that tautologies are minimally assertible, but (30b) links minimal rejectability with contradictions. The minimal assertibility (rejectability) pertains to what is unconditionally assertible (rejectable), that is, on purely logical grounds. Later (see Chap. 5, Sect. 5.4), this idea will be more precisely explained by assertibility (rejectability) on the basis of the empty set of premises.

(**DG15**) Theoretically speaking, logic can be based on falsehood as the designated logical value. Dual logic is an exact mirror of normal logic, that is, the standard logic based on the concept of truth-preserving and truh-assertibility. It is a question, why humans choose truth as the basic semantic concept. The explanation is (personal communication by Jerzy Perzanowski) that truth, contrary to falsehood, reveals facts, and we, human being, are more frequently involved in facts that their absence. An additional factor is that logic based on truth accumulates information, while dual logic disperses it. Conjunction is a connective that plays an important role in accumulating information. If $A \wedge B$ is assertible, then A and B are assertible too; the reverse link holds as well. The symmetry is broken in dual logic, because if $A \wedge B$ is rejectable, I cannot say which of the pair $\{A, B\}$ contributes to this situation or, in other words, is responsible that $A \wedge B$ can be rejected. In general, normal (that is, truth preserving) logic transmits information smoothly, but the dual logic contributes to its dispersion. On the other hand, since the meaning of connectives is associated with assertibility, rejection-rules are expressible in the metalanguage.

This observation may shed light on the problem of the genesis of logic. It seems to open the way to a naturalistic account of logic as related to facts.►

Another problem concerning the relation of logic and truth focuses on the formal (that is independent of the content of sentences on which the operator ‘it is true that’ acts) properties of the latter. Theories displaying such properties are called (formal) logics of truth (see Turner 1990, Turner 1990a, Von Wright 1996, Woleński 2004b, Woleński 2014a). I will discuss some very elementary issues with the help of diagram (**D1**):



Interpret the Greek letters as follows (I use the same letters for the set of true sentences and truth-predicates; the same applies to falsehood):

α as A is true ($\mathbf{T}(A)$);
β as $\neg A$ is true ($\mathbf{T}(\neg A)$);
γ as $\neg(\neg A$ is true$)$ ($\neg\mathbf{T}(\neg A)$);
δ as $\neg(A$ is true$)$ ($\neg\mathbf{T}(A)$);
κ as A
λ as $\neg A$
ε as A is true or $\neg A$ is true ($\alpha \vee \beta$; $\mathbf{T}(A) \vee \mathbf{T}(\neg A)$);
ζ as $\neg(\neg A$ is true$) \wedge (\neg A$ is true$)$ ($\gamma \wedge \delta$; $\neg\mathbf{T}(\neg A) \wedge \mathbf{T}(\neg A)$.

Diagram (**D1**) is a generalization (see Woleński 2007) of the well-known square of opposition for categorical sentences (‘Every *S* is *P*’, ‘No *S* is *P*’, ‘Some *S* are P’, ‘Some *S* are not *P*’) and interprets truth as a modal concept (one from the variety of so-called althetic modalities). Following, the traditional principles of opposition, w have several logical relationships summarized in

(31) (a) $\alpha \Rightarrow \varepsilon$; $\mathbf{T}(A) \Rightarrow (\mathbf{T}(A) \vee \mathbf{T}(\neg A))$;
(b) $\alpha \Rightarrow \gamma$; $\mathbf{T}(A) \Rightarrow \neg\mathbf{T}(\neg A)$;
(c) $\beta \Rightarrow \varepsilon$; $\mathbf{T}(\neg A) \Rightarrow \mathbf{T}(\neg A) \vee \mathbf{T}(A)$;
(d) $\beta \Rightarrow \delta$; $\mathbf{T}(\neg A) \Rightarrow \neg\mathbf{T}(A)$;
(e) $\neg(\alpha \wedge \beta)$; $\neg(\mathbf{T}(A) \wedge \mathbf{T}(\neg A))$;
(f) $\alpha \Rightarrow \kappa$; $\mathbf{T}(A) \Rightarrow A$;

(g) $\beta \Rightarrow \lambda$; $\mathbf{T}(\neg A) \Rightarrow \neg A$;
(h) $\neg(\kappa \wedge \lambda)$; $\neg(A \wedge \neg A)$;
(i) $\kappa \vee \lambda$; $A \vee \neg A$;
(j) $\gamma \vee \delta$; $\neg\mathbf{T}(\neg A) \vee \neg\mathbf{T}(A)$;
(k) $\neg(\alpha \Leftrightarrow \delta)$; $\neg(\mathbf{T}(A) \Leftrightarrow \neg\mathbf{T}(A))$;
(l) $\neg(\beta \Leftrightarrow \gamma)$; $\neg(\mathbf{T}(\neg A) \Leftrightarrow \neg\mathbf{T}(\neg A))$;
(m) $\zeta \Rightarrow \gamma$; $\neg\mathbf{T}(\neg A) \wedge \neg\mathbf{T}(A) \Rightarrow \neg\mathbf{T}(\neg A)$;
(n) $\zeta \Rightarrow \delta$; $\neg\mathbf{T}(\neg A) \wedge \neg\mathbf{T}\ (A) \Rightarrow \neg\mathbf{T}(A)$;
(o) $\neg(\varepsilon \Leftrightarrow \zeta)$; $\neg((\mathbf{T}(A) \vee \mathbf{T}(\neg A)) \Leftrightarrow (\neg\mathbf{T}(\neg A) \wedge \neg\mathbf{T}(A)))$;
(p) $\alpha \vee \beta \vee \zeta$; $\mathbf{T}(A) \vee \mathbf{T}(\neg A) \vee \neg\mathbf{T}(\neg A) \wedge \neg\mathbf{T}(A)$;
(q) $\varepsilon \vee \zeta$; $(\mathbf{T}(A) \vee \mathbf{T}(\neg A)) \vee (\neg(\mathbf{T}(\neg A) \wedge \neg\mathbf{T}(A))$.

The points in (31) are justified by tautologies of classical logic plus relations derived from the principles of the generalized logical square. In other setting, we have (I list only some cases):

(32) (a) α entails γ (in the traditional vocabulary: γ is subordinated to α) (31b);
(b) β entails δ (31d);
(c) α entails κ (31f);
(d) α and β are contraries, that is conjunction is always false (31e);
(e) γ and δ are subcontraries, that is, their disjunction is always true (31f);
(f) α and δ are contradictories, that is, if one is true, the other is false; the same holds for β and γ, and κ and λ (31f), (31g), (31h), (31i);
(g) the disjunction of α, β and ζ exhausts all possible cases (31n).

Since (31a)–(31o) are theorems, they can be prefixed by the universal quantifier. For example, (31n) becomes

(33) $\forall A(\mathbf{T}(A) \vee \mathbf{T}(\neg A) \vee \neg\mathbf{T}(\neg A) \wedge \neg\mathbf{T}(A))$.

Now interpret β as 'A is false' ($\mathbf{F}(A)$). Since ε is not a theorem of the logic of diagram (**D1**) (**D1**-logic), that is, a logic which generates the principles related to α–ζ, the formula

(34) $\forall A(\mathbf{T}(A) \vee \mathbf{F}(A))$

is not a theorem. This means that (**BI**) does not hold by the way of logical necessity or absolutely unconditionally. Yet (**D1**)-logic becomes an extension of classical propositional logic by adding the operators **T**, **F** and (**BI**). Incidentally, the above analysis shows that Frege's view about logic as explicating the laws of truth is too simplified, because one can say that logic is based on the principles of falsehood.

Since (34) is not logically valid, it can be negated without producing a contradiction. The related negation leads to

(35) $\exists A(\neg\mathbf{T}(\neg A) \wedge \neg\mathbf{T}(A))$,

which is equivalent (due to the definition of $\mathbf{F}(A)$) to

(36) $\exists A(\neg\mathbf{T}(A) \wedge \neg\mathbf{T}(A))$.

This last formula admits sentences that are neither true nor false. This opens the way for many-valued logic or logic with truth-value gaps.

(**DG16**) Assume that the formula ζ is universally true. This means that no sentence is true or false, that is, $\neg\exists A(\mathbf{T}(A) \vee \mathbf{F}(A))$ or, equivalently and provided that $\neg\mathbf{T}(A) \Leftrightarrow \mathbf{T}(\neg A)$ and $\neg\mathbf{F}(A) \Leftrightarrow \mathbf{F}(\neg A)$, we obtain the formula $\forall A(\neg\mathbf{T}(A) \wedge \neg\mathbf{F}(A))$. Consequently, the initial assumption that ζ is universally true, is not logically false, although it is inconsistent with (30), that is, with the constraint of minimal assertibility (something has to asserted). Saying that tautologies and, *a* fortiori, counter-tautologies lack truth-values is perhaps inconvenient, but logically possible, although at odds with a natural assumption that logical truths deserves to be asserted, but logical falsehoods—rejected. As a matter fact, Wittgenstein defended this view in his early work. He maintained that tautologies are senseless, that is, neither true nor false (see also below), but it is still debatable whether his theory admits that we are cognitively neutral to counter-tautologies.►

If (34), that is (**BI**), is added as a truth-principle, the diagram (**D**1) is reduced to its segment α–β. In particular, we then have the equivalence

(37) $\mathbf{T}(A) \Leftrightarrow \neg\mathbf{T}(\neg A)$,

which can be considered as another version of (**BI**). Decomposition of (37) (via propositional calculus) gives

(38) $\mathbf{T}(A) \wedge \neg\mathbf{T}(\neg A) \vee \neg\mathbf{T}(A) \wedge \mathbf{T}(\neg A)$.

Applying the definition of $\mathbf{F}(A)$, we obtain

(39) $\mathbf{T}(A) \wedge \neg\mathbf{F}(A) \vee \neg\mathbf{T}(A) \wedge \mathbf{F}(A)$.

Since $\neg\mathbf{F}(A)$ is equivalent to $\neg\mathbf{T}(\neg A)$, (39) also expresses (**BI**). This reasoning is another way to demonstrate that the principle of bivalence, contrary to many traditional views on the called the highest principles of thinking, does not hold with logical necessity.

Now I consider what happens under the following interpretation:

α as 'it is necessary that A'; $\Box(A)$;
β as 'it is impossible that A'; $\Box(\neg A)$;
γ as 'it is possible that A'; $\Diamond(A)$;
δ as 'it is possible that $\neg A$'; $\Diamond(\neg A)$;
κ as $\mathbf{T}(A)$;
λ as $\mathbf{T}(\neg A)$;
ε as 'it is necessary that A or it is impossible that A'; $\Box(A) \vee \Box(\neg A)$;
ζ as 'it is possible that A and it is possible $\Diamond$(or it is contingent that A).

The relations between α, β, γ and δ remain exactly the same as in (31). In addition, we easily calculate

(40) (a) $\alpha \Rightarrow \kappa$; $\Box(A) \Rightarrow \mathbf{T}(A)$;
(b) $\beta \Rightarrow \lambda$; $\Box(\neg A) \Rightarrow \mathbf{T}(\neg A)$;
(c) $\kappa \Rightarrow \gamma$; $\mathbf{T}(A) \Rightarrow \Diamond(A)$;
(d) $\lambda \Rightarrow \delta$; $\mathbf{T}(\neg A) \Rightarrow \Diamond(\neg A)$;
(e) $\neg(\kappa \Leftrightarrow \lambda)$; $\neg(\mathbf{T}(A) \Leftrightarrow \mathbf{T}(\neg A))$.

(40e) minus the definition of $\mathbf{F}(A)$ as $\mathbf{T}(\neg A)$ is demonstrably weaker then (**BI**), because it is reduced to

(41) $\mathbf{T}(\neg A) \vee \neg\mathbf{T}(\neg A)$.

Even if (41) is considered as a version of the principle of excluded middle, because it falls under the scheme $A \vee \neg A$ taken as general, it is still weaker than (**MEM**) and acceptable in logics of truth without bivalence. Using $\mathbf{F}(A)$ in place of $\mathbf{T}(\neg A)$ reduces **D1**-logic to the classical propositional system. Diagram (**D1**) shows the results of introducing the concept of necessary truth and truth *simpliciter*. If so, one can argue that necessary truths and truths can be distinguished, at least from a logical point of view.

The universal generalization of ε, that is:

(42) $\forall A(\Box(A) \vee \Box(\neg A))$.

This formula codes the statement that every truth is necessary or impossible (briefly, but not quite accurately: every truth is necessary, because if A is impossible, then $\neg A$ is necessary). On the other hand, the formula:

(43) $\forall A(\Diamond(A) \wedge \Diamond(\neg A))$

displays the opinion that every truth is contingent. If one says that what is represented in the diagram (**D1**), governs our thinking about truths in the sense that we have necessary and contingent truths, then the statement 'A is true' is ambiguous, because it can mean either that 'A is necessary' or 'A is possible'. On the other hand, γ and δ play rather an auxiliary role in this framework, and serve as devices to define the concept of accidental truth (I opt for understanding of contingency as accidentality, not not-contingency). Assume that A is a possible truth, that is, $\Diamond(A)$ holds. If A is also necessary, which is possible as was seen in (31b), further consideration is not needed. Thus, it remains to review what happens when A is possible, but not necessary. Under this assumption, A cannot be impossible. Thus, A is either possible and true or possible and false (it does not matter whether A is false or $\neg A$ true). The latter case immediately implies the contingency of A, because we have $\mathbf{T}(\neg A)$ implying $\Diamond(\neg A)$. In the former case, if A is not necessary, its negation is possible and it too is contingent. These facts will be essentially used in Chap. 9, Sect. 9.5 in the analysis of logical determinism.

If we return to the first interpretation of (**D1**), the formula

(44) $\mathbf{T}(A) \Rightarrow A$

is of a special interest. It holds due to logic, but its converse, that is

(45) $A \Rightarrow \mathbf{T}(A)$

has no justification in the **D1**-logic, because if we admit many-valueness or truth-value gaps, the interpretation of the antecedent of (44) cannot be the same as the interpretation of the consequent in (43). Assume that A is valued as *neutrum* as in Łukasiewicz's three-valued logic. Of course, it does not invalidate (45), because in this case both its ingredients are false. Now, if A is an example of *neutrum*, it is true that it is a *neutrum*. On the other hand, it is false that A is true in this case. This observation is very important, because the conjunction of (43) and (44) gives

(46) $\mathbf{T}(A) \Leftrightarrow A$,

This is another version of **T**-scheme (we can read this as it is true that A if and only if A) that plays a crucial role in many truth-theories, including **STT** (see Chap. 3, Sect. 3.7, Chap. 7, Sects. 3.2, 3.3, 3.4 and Chap. 9, Sect. 9.3). The fact that only a half of (47) has a purely logical justification via the logic of truth can be regarded as an informal demonstration that **T**-equivalences are not tautologies (see more in Chap. 9, Sect. 9.3), but logical contingencies (accidentalities). If (46) is accepted, the points κ and λ simply disappear from the analyzed conceptual scheme, but it does not affect γ and δ, and they remain as before. This means that (**BI**) is stronger than **T**-scheme. Although it is possible to introduce bivalence by (24), this move makes the concept of non-truth ambiguous, because it then means either being false, or having the other logical value, or representing a truth-value gap. Hence, (**BI**) in the traditional version is not trivial. Moreover, one can retain the **T**-scheme even if (**BI**) is dropped, provided, of course, that **T**-equivalences hold only for true sentences. In fact, it is easy to prove that (**BI**) entails **T**-scheme. If so, this means that we must infer the formula $\mathbf{T}(A) \Leftrightarrow A$ from the disjunction $\mathbf{T}(A) \vee \mathbf{F}(A)$. Assume that $\mathbf{T}(A)$ holds. If $\mathbf{T}(A)$ is true, A is also true. This gives $\mathbf{T}(A) \wedge A$, and $\mathbf{T}(A) \Rightarrow A$, as well as $A \Rightarrow \mathbf{T}(A)$. These both conditionals demonstrate the equivalence $\mathbf{T}(A) \Leftrightarrow A$. Furthermore, if $\mathbf{F}(A)$ is true, $\neg\mathbf{T}(A)$ and $\neg A$ are true. This leads to the formulas $\neg\mathbf{T}(A) \Rightarrow \neg A$ and $\neg A \Rightarrow \neg\mathbf{T}(A)$. Contrapositions of both implications result in $\mathbf{T}(A) \Leftrightarrow A$. Thus, **T**-scheme is proved from (**BI**). On the other hand, the assertion that **T**-scheme holds for true sentences (in fact, it also holds for falsehoods via contraposition), does not entail that (**BI**) is valid (universally true).

Our diagram does not generate all of the principles for **T**-logic (the logic of truth). If one accepts (see Turner 1990, p. 25) the following formulation for the classical case (note that I focus only on the propositional part and omit the rules of inference):

(**TA1**) $\mathbf{T}(A) \Leftrightarrow A$, for all atomic A;
(**TA2**) $\mathbf{T}(A \wedge B) \Leftrightarrow \mathbf{T}(A) \wedge \mathbf{T}(B)$;
(**TA3**) $\mathbf{T}(\neg A) \Leftrightarrow \neg\mathbf{T}(A)$;

(**TA4**) $\mathbf{F}(A) \Leftrightarrow \mathbf{T}(\neg A)$;
(**TA5**) $\neg(\mathbf{T}(A) \wedge \mathbf{F}(A)$,

This **T**-logic extends **D1**-logic to a system in which all **T**-biconditionals become theorems.

(**DG17**) An additional restriction must be made in order to block inconsistency of **T**-logic. This is possible by excluding the paradoxical sentence in advance, for example, by prohibiting self referential contexts (see Chap. 6). However, it is interesting that if we reject (44) the Liar Paradox (**LP** for brevity) disappears (see Turner 1990, p. 24, and Halbach 2011, Chaps. 13 and 15).►

(TA2) can be replaced by:

(47) $\mathbf{T}(A \Rightarrow B) \Rightarrow (\mathbf{T}(A) \Rightarrow \mathbf{T}(B))$.

This last formula means that truth is monotonic (distributive over implication), at least in the described **T**-logics.

Let τ be an arbitrary tautology and A an arbitrary contingent sentences. Since tautology is implied by everything, we have

(48) $A \Rightarrow \tau$.

Applying (47) we obtain

(49) (a) $\mathbf{T}(A \Rightarrow \tau) \Rightarrow (\mathbf{T}(A) \Rightarrow \mathbf{T}(\tau))$.
(b) $\mathbf{T}(\neg A \Rightarrow \tau) \Rightarrow (\mathbf{T}(\neg A) \Rightarrow \mathbf{T}(\tau))$.

According to Wittgenstein, if a sentence A is contingent, it is meaningful as well. Thus, A is true or false (according to the standard definition that a sentence is logically meaningful, provided that it possesses one of logical values, namely truth or falsehood; in this second case, $\neg A$ is true. A simple argument shows that τ is true independently whether A is true or false. Thus, if something is true, tautologies are true too. Wittgenstein's view (see (**DG16**)) was not correct, at least if the **T**-logic outlined above holds. Consequently, there is not very much wisdom in considering tautologies as neither true nor false (in fact, that they are true remains the only reasonable possibility.

Another system of **T**-logic was proposed by Von Wright (see Von Wright 1996). He proposes the following axioms for the basic truth-logic **CS** (I omit the rules of inference also in this case):

(**TA'1**) All tautologies of classical propositional calculus interpreted (via substitution) by formulas of the type $\mathbf{T}(A)$, their negations, conjunctions, etc.

(**TA'2**) $\mathbf{T}(A) \Leftrightarrow \mathbf{T}(\neg\neg A)$;
(**TA'3**) $\mathbf{T}(A \wedge B) \Leftrightarrow \mathbf{T}(A) \wedge \mathbf{T}(B)$;
(**TA'4**) $\mathbf{T}\neg(A \wedge B) \Leftrightarrow \mathbf{T}(\neg A) \vee \mathbf{T}(\neg B)$.

This logic has the weak rule of excluded middle in the form

(50) $\mathbf{T}(A) \vee \neg\mathbf{T}(A)$,

which follows from the formula $A \vee \neg A$. The system **CS** can be supplemented by adding (37) to its axioms. The result is that non-truth and falsehood are not distinguishable (the strong excluded middle is valid) by this addition. A further extension arises when (43)—that is, the obvious part ($\mathbf{T}$(A) $\Rightarrow A$) of **T**-scheme is added as a new principle. These brief remarks show how diagram (**D1**) is related to axiomatic approaches to truth-logic. Anyway, even if the formal properties of truth do not exhaust of the content of its concept), they contribute essentially to the understanding this notion.

The logic of truth as displayed by (**D1**) has an interesting application for analysis a celebrated Hume's thesis that is-sentences do not imply ought-sentences. The original wording of Hume's thesis is as follows (D. Hume, *A Treatise on Human Nature*, Clarendon Press, Oxford 1951, p. 469):

> I cannot forbear adding to these reasoning an observation, which may, perhaps, be found of some importance. In every system of morality, which I have hitherto met with, I have always remark'd, that the author proceeds for some time in the ordinary way of reasoning, and establishes the being of a God, or makes observations concerning human affairs; when of a sudden I am surpriz'd to find, that instead of the usual copulations of propositions, *is*, and *is not*, I meet with no proposition that is not connected with an *ought*, or an *ought not*. This change is imperceptible; but is, however, of the last consequence. For as this *ought*, or *ought not*, expresses some new relation or affirmation, tis necessary that it shou'd be observ'd and explain'd; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it.

The Hume thesis can be justified by deontic logic (see Woleński 2006, Woleński 2007). Let us return to diagram (**D1**) stipulating that (I mention only relevant points; the rest is easily reconstructed):

(a) $\Box(A)$ means 'it ought be (it is obligatory) that A;
(b) $\Diamond(A)$ means 'it is permitted that A';
(c) κ is interpreted as $\mathbf{TR}(A)$.
(d) the relations between α, β, γ and δ remain as before.

Originally, Hume's observations concerned the relation of κ to α. The Hume thesis says that κ does not imply α. Let us call this the simple Hume thesis for obligation. Implicitly, Hume also stated another thesis, namely, that α does not entail κ. Let us call this the reverse Hume thesis for obligation. However, similar principles can be formulated for permission. The simple Hume thesis for permission says that κ does not imply γ; the reverse Hume thesis for permission says that γ does not imply κ.

The exact formulation of particular Hume's theses is as follows:

(51) (a) $\neg \vdash \mathbf{T}(A) \Rightarrow \Box(A)$ (the simple Hume thesis for obligation);

(b) $\neg \vdash \Box(A) \Rightarrow \mathbf{T}(A)$ (the reverse Hume thesis for obligation);
(c) $\neg \vdash \mathbf{T}(A) \Rightarrow \Diamond(A)$ (the simple Hume thesis for permission);

This analysis shows that the Hume thesis in its original formulation is incomplete.

(**DG18**) Diagram (**D1**) in its former interpretation validates the alethic counterparts of opposites of (51b) and (51c)—that is, the formulas $\Box(A) \Rightarrow \mathbf{T}(A)$ and $\mathbf{T}(A) \Rightarrow \Diamond(A)$. Adding (44) to the **D1**-logic is innocent, but $\mathbf{T}(A) \Rightarrow \Box(A)$ (truth implies necessity or every truth is necessary) is very controversial. Thus, not all possible interpretations of the diagram (**D1**) are formally equivalent. For instance, the principle $A \Rightarrow \Box(A)$ is not admissible if the box means 'it is known that' (epistemic interpretation). The reverse implication, that is, $\Box(A) \Rightarrow A$, is frequently adopted in epistemic logic, but it is rather justified by the definition of knowledge than by a clear logical principle.►

Whether a given formula is valid (logically true) or not depends on semantics. Any serious investigation of the Hume theses requires an appeal to deontic semantics. I will use the semantics for standard deontic logic. Let as assume that we have the ordered triple (Kripke frame) $\mathbf{S} = <\mathbf{K}, \mathbf{W}^*, \mathbf{R}>$, where **K** is a non-empty set of items called possible worlds, **W*** is a designated element of **K**, usually interpreted as the real world, and **R** is a binary relation defined on **K** (the accessibility or alternativeness relation; if **W′RW**, we say that the former is a deontic alternative for the latter). **S** is a deontic frame if and only if **R** is not reflexive, that is, it is not generally true that **WRW**. In particular, we assume that if it is not the case that **W*RW*** (the real world is not a deontic alternative of itself, that is):

(52) $\Box(A)$ is true in the world **W*** if and only if $\mathbf{T}(A)$ is true in every world **W** such that **WRW***.

Intuitively, the sentence 'it is obligatory that *A*' is true in the real world **W*** if and only if *A* is true in every world **W** being a deontic alternative to **W***, that is, in a world in which all obligations valid in the real world are satisfied. Accordingly, we have a condition for permission:

(53) $\Diamond(A)$ is true in the world **W*** if and only if there is a world **W** such that **WRW***

and $\mathbf{T}(A)$ is true in **W**.

The non-reflexivity of **R** is justified by the fact that not all obligations are satisfied in the real world. If **R** is not reflexive, $\Box(A)$ can be true, but $\mathbf{TR}(A)$ false. Thus, (49b) fails as a tautology of deontic logic. The simple Hume thesis is justified, because $\mathbf{TR}(A)$ can be true (people may behave in the way described by sentence *A*), but this kind of action is not obligatory. It is also clear that $\mathbf{TR}(A)$ can be true, but *A* still not be permitted; thus (49c) is not a logical principle. As far as *A* being permitted (see 49d), this fact is consistent with the falsehood of $\mathbf{TR}(A)$. Note, finally, that deontic sentences are declaratives, not imperatives or norms (as linguistic utterances). The relation of the former to the latter comprises a separate question. Note than in Hume's quoted passage (see above) we find "*ought*, or *ought not*, expresses some new relation or affirmation" and this phrase clearly indicates that it concerns indicatives. It is important because the is/ought problem, for instance, the underivability of what ought to be from what is, appears as independent from the question whether normative utterance have the status of indicatives or are reducible to imperatives.

(**DG18**) (Inspired by some remarks of the referee). I would like to stress once again that this chapter is selective at least in two directions. Firstly, there are many questions concerning tasks of truth-theories, not discussed above—some of them were mentioned in (**DG2**). Secondly, such topics as the nature of propositions, judgements, beliefs, sentences (the token/type distinction, for instance), concepts or assertions, furthermore, various issues, partly controversial, concerning, truth-criteria and their relation to truth-definitions, existing and possible classifications of truth-theories, a closer characterization of particular truth-theories, many-valued logics, etc. could be discussed much more extensively than I did. I have two excuses. Firstly, I subordinated my considerations to the aim of later presentation of **STT**. I am aware that perhaps some questions could be discussed more widely, other (truth as a modality)—less extensively, but I also followed my subjective preferences, in particular, that, I see no chance to offer a satisfactory theory of propositions or that the logic of truth deserves more attention that it usually acknowledged. Secondly, the recent compendium of the philosophy of truth, that is Glanzberg 2018, has 832 pp., and, at least I guess so, some reviewers will have reservations about its completeness. Unfortunately, this collection appeared after my book was practically completed and could not address my remarks to the *Oxford Handbook of Truth*. I only remark that Ray 2018, a paper on **STT** is rather poor, because it practically reduces this theory to **T**-scheme and neglects most important technical as well philosophical issues (see Chaps. 7, 8 and 9).►

References

Ajdukiewicz, K. (1949). *Zagadnienia i kierunki filozofii*. Warszawa: Czytelnik; Eng. tr. *Problems & Theories of Philosophy*. Cambridge: Cambridge University Press.

Albuquerque, D. (1995). *Truth and Action in Speech acts. A Study on J. L. Austin's Essay on Truth*. New Delhi: Intercultural Publications.

Auxier, R. E., Hahn, L. E. (Eds.) (2006). *The Philosophy of Jaakko Hintikka*. La Salle: Open Court

Baldwin, T. (1991). Can there be a substantive account of truth? In Cooper, Engel (1991), 22–39.

Bennet, B. M., Hoffman, D. D., Prakash, Ch. (1989). *Observer Mechanics. A Formal Theory of Perception*. New York: Academic Press.

Beziau, J.-Y., Carnielli, W., Gabbay, D. (2007). *Handbook of Paraconsistency*. London: College Publications.

Blackburn, S., Simmons, K. (Eds.) (1999). *Truth*. Oxford: Oxford University Press.

Carnap, R. (1947). *Meaning and Necessity. A Study in Semantics of Modal Logic*. Chicago: University of Chicago Press; 2nd ed., 1956.

Cartwright, E. (1987). *Philosophical Essays*. Cambridge, Mass.: The MIT Press.

Cartwright, R. (1987a). A neglected theory of truth. In Cartwright (1987), 71–93.

Changeux, J.-P. (2002). *Physiology of Truth. Neuroscience and Human Knowledge*. Cambridge, Mass.: The Belknap Press.

Cooper, N., Engel, P. (Eds.) (1991). *New Inquiries into Meaning and Truth*. New York: St. Martin's Press.

David, M. (2004). Theories of truth. In Niiniluoto, Sintonen, Woleński, 331–414.

Davidson, D. (1996). The folly to trying to define truth. *Dialogue and Universalism*, 6, 39–53.

Dumitriu, M., Sandu, G. (Eds.) (2013). *Truth*. Bucureşti: Editura Universitãtii din Bucureşti.

Dummett, M. (1999). Of what kind of things Is truth a property? In Blackburn, Simmons (1999), 264–281.

Edwards, D. (2018). *The Metaphysics of Truth*. Oxford: Oxford University Press.

Engel, P. (1991). *The Norm of Truth: An Introduction to the Philosophy of Logic*. Toronto: University of Toronto Press.

Gaskin, R. (2013). *Language, Truth, and Literature. A Defence of Literary Humanism.* Oxford: Oxford University Press.

Glanzberg, M. (Ed.) (2018). *The Oxford Handbook of Truth*. Oxford: Oxford University Press.

Haack, S. (1978). *Philosophy of Logics*. Cambridge: Cambridge Univeristy Press.

Halbach, V. (2011). *Axiomatic Theories of Truth*. Cambridge: Cambridge University Press.

Halbach, V., Horsten, L. (Eds.) (2002). *Principles of Truth*. Frankfurt am Main: Dr. Hänsel-Hohenhausen.

Halldén, S. (1960). *True Love, True Humour and True Religion*. Lund: CWK Gleerup.

Hempel, K. (1935). On the logical positivistic theory of truth. *Analysis*, 2, 49–55; repr. in Horwich (1994), 80–89.

Hofstadter, A. (1965). *Truth and Art*. New York: Columbia University Press.

Horwich, P. (Ed.) (1994). *Theories of Truth*. Aldershot: Dartmouth.

Huby, P., Neal, G. (Eds.) (1989). *The Criterion of Truth*. Liverpool: Liverpool University Press.

Jago, M. (2018). *What is Truth*. Oxford: Oxford University Press.

Johnson, E. L. (1992). *Focusing on Truth*. London: Routledge.

Kastil, A. (1947). Wahrheit und Sein. *Zeitschrift für philosophische Forschung*, I, 483–498.

Kirkham, R. L. (1992). *Theories of Truth. A Critical Introduction*. Cambridge, Mass.: The MIT Press.

Kolak, D., Symons, J. (Eds.) (2004). *Quantifiers, Questions and Quantum Mechanics*. Dodrecht: Springer.

Künne, W. (1985). Wahrheit. In Martens, Schnädelbach (1985), 116–171.

Künne, W. (2003). *Conceptions of Truth*. Oxford: Clarendon Press.

Linke, P. (1949). Die mehrwertige Logiken und Wahrheitsproblem. *Zeitschrift für philosophische Forschung*, III, 378–398, 530–546.

Malinowski, G. (1994). *Many-Valued Logics*. Oxford: Oxford University Press.

Marsonet, M., Benzi, M. (Eds.) (2004). *The Proceedings of Genoa Conference on Logic and Metaphysics*. Name: Fiesola.

Martens, E., Schnädelbach, H. (1985). *Philosophie. Ein Grundkurs*. Reimbek bei Hamburg: Rowohlt.

McCloskey, H. J. (1969). *Meta-Metaethics and Normative Ethics*. Berlin: Springer.

Moore, G. E. (1899). The nature of judgement. *Mind*, 5, 176–193; repr. in G. E. Moore (1993), 1–19.

Moore, G. E. (1993). *Selected Writings*. London: Routledge.

Niiniluoto, I., Sintonen, M., Woleński, J. (Eds.) (2004). *Handbook of Epistemology*. Dodrecht: Kluwer.

Patterson, D. M. (1999). *Law and Truth*. Oxford: Oxford University Press.

Popper, K. (1955). A note on Tarski's definition of truth. *Mind*, 64, 388–391; rept. Popper (1972), 335–340.

Popper, K. (1972). *Objective Knowledge. An Evolutionary Approach*. Oxford: Clarendon Press (repr. with additions, 1974).

Priest, G. (2008). *An Introduction to Non-Classical logic*. Cambridge: Cambridge University Press.

Puntel, L. B. (1990). *Grundlagen einer Thorie der Wahrheit*. Berlin: de Gruyter.

Ray, G. (2018). Tarski on the concept of truth. In Glanzberg (2018), 695–719.

Rescher, N. (1969). *Many-Valued Logic*. New York: McGraw-Hill.
Ritter, J., Gründer, K. H., Gabriel, G. (Eds.) (1971–2007), *Historischer Wörterbuch der Philosohischen Begriffen*, v. 1–13. Darmstadt: Wissenschafliche Buchgessellschaft.
Russell, B. (1903). *The Principles of Mathematics*. Cambridge: Cambridge University Press.
Sher, G. (1999). On the possibility of a substantive theory of truth. *Synthese*, 117, 133–172.
Soames, S. (1999). *Understanding Truth*. Oxford: Oxford University Press.
Tarski, A. (1933). *Pojęcie prawdy w językach nauk dedukcyjnych* (The Concept of Truth in Languages of Deductive Sciences). Warszawa: Towarzystwo Naukowe Warszawskie; Germ. tr. (with additions) as Tarski (1935).
Tarski, A. (1935). *Der Wahrheitsbegriff in den formalisierten Sprachen*. Studia Philosophica, 1, 261–405; repr. in A. Tarski (1986), v. 2, 51–198; Eng. tr., *The Concept of Truth in Formalized Languages*, in Tarski (1956), 152–278 (page-references Tarski (1933)).
Tarski, A. (1956). *Logic, Semantics, Metamathematics. Papers of 1923 to 1938* (tr. by J. H. Woodger,). Oxford: Clarendon Press; 2nd ed., Indianapolis: Hackett Publishing Company 1983.
Tarski, A. (1986). *Collected Papers*, v. 1–4. Basel: Birkhäuser; 2nd ed., 2018.
Turner, R. (1990). *Truth and Modality for Knowledge Representation*. London: Pitman.
Turner, R. (1999). Logics of truth. *Notre Dame Journal of Formal Logic, 31*, 308–329.
Twardowski, K. (1900). O tzw. prawdach względnych (On so-called relative truths). In *Księga Pamiątkowa Uniweysytetu Lwowskiego ku uczczeniu pięćsetnej rocznicy Fundacji Jagiellońskiej Uniwersytetu Krakowskiego* (Memorial Book of Lvov University on the Occasion of 500thAnniversary of Jagiellonian Fund—The University of Kraków), 64–93. Lwów: Nakładem Uniwersytetu Lwowskiego, 64–93; Eng. tr. in Twardowski (1999), 147–169.
Twardowski, K. (1912). O czynnościach i wytworach. Kilka uwag z pogranicza psychologii, gramatyki i logiki (Actions and products. remarks from the boderline of psychology, grammar and logic). In *Księga Pamiątkowo ku uczczeniu 250-tej rocznicy założenia Uniwersytetu Lwowskiego przez króla Jana Kazimierza* (Memorial Book on the Occasion of Establishment of Lvov University by the King Jan Kazimierz), v. II. Lwów: Nakładem Uniwersytetu Lwowskiego, 1–33; Eng. tr. in Twardowski (1999), 103–132.
Twardowski, K. (1999). *On Actions, Products and Other Topics in Philosophy* (A. Szylewicz, Trans.). Amsterdam: Rodopi.
Van Fraassen, B. (1971). *Formal Semantics and Logic*. New York: Macmillan.
Van Zantwijk, T., Gabriel G. (2001). Urteil. In Ritter, Gründer, Gabriel (Eds.) (1971–2007), v. XI, 430–458.
Von Wright, G. H. (1996). *Six Essays in Philosophical Logic*. Helsinki: Societas Philosophica Fennica.
Von Wright, G. H. (1996a). Truth logics. In Von Wright (1996), 71–91.
Vroom, H. M. (1989). *Religions and Truth: Philosophical Reflections*. Amsterdam: Rodopi.
Wansing, H., Shramko Y. (2011). *Truth and Falsehood: An Inquiry into Generalized Logical Values*. Berlin: Springer.
Weaver, N. (2015). *Truth & Assertibility*. Singapore: World Scientific.
Weingartner, P. (2000). *Basic Questions on Truth*. Dordrecht: Kluwer.
Wentscher, E. (1941). *Relative oder absolute Wahrheit? Eine Studie*. München: Reinhardt.
Williams, B. (2004). *Truth and Truthfulness. An Essay in Genealogy*. Princeton: Princeton University Press.
Williams, C. J. F. (1976). *What is Truth?* Cambridge: Cambridge University Press.
Woleński, J. (1989). *Logic and Philosophy in the Lvov-Warsaw Lchool*. Dordrecht: Kluwer.
Woleński, J. (1995). Logic and falsity. *Logica Yearbook* 1994, 95–105.
Woleński, J. (2004b). On some formal properties of truth. In Kolak, Symons (2004), 195–207; repr. in Woleński (2011), 91–100.
Woleński, J. (2004c). What are truth-bearers? *Reports on Philosophy*, 22, 34–47; repr. in Woleński (2011), 145–154.
Woleński, J. (2004d). Is Truth an ontological concept? In Marsonet, Benzi (2004), 133–139.

Woleński, J. (2006). A generalization of Hume's thesis. *Logica Yearbook 2005*, 215–223; repr. in Woleński (2011), 155–162.

Woleński, J. (2006a). Tarskian and post–Tarskian truth. In Auxier, Hahn (2006), 647–672: repr. in Woleński (2011), 163–182.

Woleński, J. (2007). Applications of squares of opposition and their generalizations in philosophical analysis. *Logica Universalis*, 2(1), 13–29; repr. Woleński (2011), 255–269.

Woleński, J. (2011). *Essays on Logic and Its Applications in Philosophy*. Frankfurt a. M.: Peter Lang.

Woleński, J. (2013). Ens et verum convertuntur? (Are being and truth convertible. A contermporary perspective)? In Dumitriu, Sandu (2013), 75–83; repr. in Woleński (2018), 107–114.

Woleński, J. (2014a). An abstract approach to bivalence. *Logic and Logical Philosophy*, 23(1), 3–14; repr. in Woleński (2018), 115–124.

Woleński, J. (2018). *Logic and its Philosophy*. Frankfurt a. M.: Peter Lang.

Woods, J. (2018). *Truth in Fiction. Rethinking Its Logic*. Springer: Berlin.

Zinoviev, A. A. (1963). *Philosophical Problems of Many-Valued Logic*. Dordrecht: Reidel.

Chapter 5
Matters of Logic

Abstract Since **STT** is a logical theory, its relation to logic are very close. The task of this chapter consists in presenting logical concepts and theories relevant for the further discussion of **STT**. Particular sections are devoted to propositional calculus, first-order logic, metalogic, definitions of logic (the universality of logic is its essential attribute) and historical notes on metalogic and metamathematics.

5.1 Introduction

Preparing a formal apparatus for discussing **STT** constitutes the primary task of this chapter. Since I will unveil this theory in the framework of first-order logic (**FOL**, henceforth), this system is considered as the model of logic. This perspective is essentially different than that used in Tarski 1933 and its translations into German (Tarski 1935) and English (Tarski 1956). Originally, **STT** was constructed for a formal system similar to the simple theory of types (see Chap. 6, Sect. 6.3), but the first-order formalism is its contemporary setting, which is adopted in all contemporary textbooks of logic or model theory (it was introduced in Tarski, Vaught 1957; in many respects I follow Grzegorczyk 1974; see also Hodges 2004 for a more advanced treatment, in particular on conditions imposed on languages having Tarski's truth definition). The presentation of **FOL** also covers its metalogical (metamathematical) properties, because that has a particular significance for **STT**, which, to repeat a remark from the Introduction, is the foundation of model theory, one of the fundamental parts of modern logic. Note, however, that this chapter is not intended as a textbook of logic or its shortening. Although I into this fragment of the book included some topics which exceed the logical environment of **STT**, I believe that they are relevant for proper understanding of this theory as closely related to logic.

Logic was always considered as part (frequently very important) of philosophy or at least as something very closely related to philosophical enterprise, in particular as its organon (in the Aristotelian sense). Hence, the nature of logic has constantly been debated among philosophers ever since logic was born (or discovered) in

J. Woleński, *Semantics and Truth*, Logic, Epistemology, and the Unity of Science 45, https://doi.org/10.1007/978-3-030-24536-8_5

ancient Greece. These discussions focused on a considerable number of fundamental issues. Let me mention only three of them. What are properties of logic? What are the functions of logic? What is the scope of logic? These issues are interconnected in many ways. Perhaps it is convenient to begin with the distinction between logic *sensu largo* (logic in a wide sense) and logic *sensu stricto* (logic in a narrow sense, formal logic). Whereas the latter can be identified with formal logic, the former includes semantics (semiotics) and methodology of science. My remarks in this chapter concern logic *sensu stricto*. Consequently, the problem of properties of formal logic is of the utmost importance for its scope and functions. Logic *sensu stricto* studies deductive, correct, or demonstrative inferences and does that by reference to the form of arguments while abstracting from their content. Yet it is not true that formal logic has no connection with semantics, because it must elaborate how logical languages are semantically evaluated (note that Chap. 6 will discuss semantics from a more general point of view, not limited to its place in formal logic).

The second important distinction related to the nature of logic is that of *logica docens* (theoretical logic) and *logica utens* (applied logic). Constructing logical systems and investigating their properties constitute the merit of *logica docens* (including metalogic). On the other hand, almost everybody agrees that we use logic and should do that (I do not enter into details concerning this claim). Petrus Hispanus a medieval logician said that *dialectica* (that is, logic) *est art artium et scientia scientiarum ad omnium aliarum scientiarum methodorum principia viam habent* (logic is the art of all arts and science of sciences, which provides methods for all other sciences). This immediately suggests that theoretical logic is universal and universally applied. The essence of Petrus Hispanus' quoted *dictum* about logic as *scientia scientiarum* contains the most fundamental intuition about logic. Gödel (Gödel 1944, p. 125), Tarski and Quine (Quine 1970, p. 102) expressed the same idea in the following passages:

> […] [logic] is a science prior to all others, which contains the ideas and principles underlying all sciences.

> […] the word "logic" is used […] in the present book […] as the name of the discipline which analyses the meaning of the concepts common to all sciences, and establishes general laws governing these concepts.

> The lexicon is what caters distinctively to special tastes and interests. Grammar and logic are the central facilities, serving all comers.

Thus, one can think about the view of Petrus Hispanus-Gödel-Tarski-Quine as suggesting that logic is universally applicable.

The collected material suggests a distinction of three notions of the universality property of logic:

(**Un1**) logic is universal because it is universally applicable;
(**Un2**) logic is universal because it is topic-neutral;
(**Un3**) logic is universal because its principles are universally valid.

Although (**Un1**)–(**Un3**) can be attributed to *logica utens* as well as to *logica docens*, (**Un1**) seems to be primarily addressed rather to the former, but (**Un2**) and (**Un3**) viewed as features of the latter. Since *logica utens* acts as an applied science, its essence consists in formulating rules for performing inferences. On the other hand, *logica docens* has fairly descriptive tasks. It aims at a theoretical description of the world of logic, whatever this reality is or seems to be. Yet theoretical logic seems to provide devices employed in *logica utens*. One of my tasks consists in giving a formal characterization of the universality property of logic and demonstrating that (**Un1**)–(**Un3**) are equivalent.

In the history of logic, appeared also some projects of so-called *logica magna* (grand logic), systems to cover the entire science, or at least—mathematics. Leibniz's *characteristica universalis* and *calculus ratiocinator*, the systems of Frege and Russell, dictated by their logicism, or Leśniewski's logic can be taken as illustrations of this understanding of logic. Today, we have rather more moderate accounts, like this one (Barwise 1985, pp. 4–5):

> […] logic consists of a collection of mathematical structures, a collection of formal expressions, and a relation of satisfaction between the two […]. We can say, then, that a logic is something we construct to study the logic of some parts of mathematics.

Logic in this view offers various suitable tools envisaged as possible descriptions of various mathematical structures. There is no doubt that any good logic must be strongly expressive in order to capture as much mathematical content as possible. The demand of a great expressive power of logic leads to a fourth notion of logical universality (I will refer to it by (**U4**)). According to this notion of universality, to say that logic is universal means that its content is rich: the universality of logic in this sense is, so to speak, directly proportional to the content of logical theorems. How are the notions of the universality of logic displayed by (**Un1**)–(**Un4**) mutually related? Assuming that (**Un1**)–(**Un3**) are equivalent, I will argue that (**Un4**) captures a different concept of universality.

(**DG1**) What is logical correctness? Consider that a mathematician proves a statement *B* on the basis of a statement *A*, that is, demonstrates that *A* logically entails *B*. If someone were to show that this is not the case, the proof loses its validity and must be improved; otherwise, it is not acceptable. This reflects that logical correctness associated with the conditions of provability is completely determined; the correct proof is simply a proof, the incorrect proof is not a proof at all, even if it can be improved. On the other hand, if we pass to what pertains to logical analysis of concepts, the situation radically changes. For instance, logical empiricists claimed that their analysis of meaningfulness was logical and correct because it displayed the logical grammar of language, which was identified with logic. If they had in mind logic in the narrow sense, this claim was erroneous, but if they operated with logic as applied to philosophy, the correctness of their analysis of meaningfulness becomes another matter and must be very carefully justified. At this point I can return to the methodological remarks formulated in the Introduction—in particular, to the idea of interpretative consequence. The use of logical analysis in philosophy does not consist in a straightforward application of logic, but is surrounded by many extralogical claims, in particular, various conditions of adequacy. Hence, the correctness of logical analysis cannot be exhausted solely by formal criteria. This point is of the utmost importance for any philosophical analysis of **STT**, because this theory of truth is deeply rooted in a definite logical environment. Hence, to identify what is strictly logical in **STT** and what is logical in a broader sense becomes a very significant issue. One of the most confusing things in philosophy is to say that something 'is logical', or even 'purely logical', without explaining that means.►

(**DG2**) The presentation of logic in Sect. 5.2 and metalogic in Sect. 5.3 is standard, the possible exception of some (minor) points concerning the concept of interpretation in predicate calculus. I follow the material accessible from textbooks (see, Pogorzelski 1994 for a general survey of logic and metalogic as well as their basic concepts and results). Since this chapter is not intended as a substitute of a text-book of logic, I assume knowledge of elementary concepts of set theory and rudimentary techniques of logic.►

5.2 Logical Calculi

I will present so-called elementary logic, that is, propositional calculus **PC** (the name of this system does not suggests that propositions are its items; recall, that I use the terms 'sentence' and 'proposition' as equivalents, at least in the context of logic) and **FOL**. Every logical calculus **LC** is an ordered quadruple <**AL**, **L**, **AX**, **RI**>, where **AL** is an alphabet of **L**, **L** is a language (a set of well-formed formulas) of **LC**, **AX** is a set of axioms, and **RI** is a set of rules of inferences. All elements of **LC** can be indexed by reference to **PC** or **FOL**, for instance $\mathbf{L}_{\mathbf{PC}}$ or $\mathbf{L}_{\mathbf{FOL}}$.

A. Propositional Calculus: Syntax

The alphabet of **PC** ($\mathbf{AL}_{\mathbf{PC}}$) = {$\mathbf{VAR}^{\mathbf{P}} \cup \mathbf{CONST}^{\mathbf{PC}} \cup$ {(,)}}, where $\mathbf{VAR}^{\mathbf{P}}$ is a denumerably infinite list of propositional variables: $p_1, p_2, p_3, \ldots$; $\mathbf{CONST}^{\mathbf{PC}}$—the finite list of propositional logical constants (propositional connectives) $\neg$ (negation, 'not', 'it is not the case'), $\wedge$ (conjunction; 'and'), $\vee$ (disjunction, 'or' in the inclusive sense), $\Rightarrow$ (implication, if, then), $\Leftrightarrow$ (equivalence; if and only if); {(,)}—the set of parentheses (– left;)—right. The language of **PC** ($\mathbf{L}_{\mathbf{PC}}$) = the set $\mathbf{FOR}_{\mathbf{PC}}$ of well-formed formulas (wffs) of $\mathbf{L}_{\mathbf{PC}}$:

(**Df1**) $\mathbf{FOR}_{\mathbf{PC}}$ is the smallest set such that
(a) for every i (i = 1, 2, 3, …), $p_i \in \mathbf{FOR}_{\mathbf{PC}}$;
(b) if $A, B \in \mathbf{FOR}_{\mathbf{PC}}$, then $\neg A \in \mathbf{FOR}_{\mathbf{PC}}$, $A \wedge B \in \mathbf{FOR}_{\mathbf{PC}}$, $A \vee B \in \mathbf{FOR}_{\mathbf{PC}}$, $A \Rightarrow B \in \mathbf{FOR}_{\mathbf{PC}}$, $A \Leftrightarrow B \in \mathbf{FOR}_{\mathbf{PC}}$.

(**DG3**) Definition (**Df1**) is an example of an inductive definition. The term 'smallest set' refers to the product of the all sets satisfying the prescribed constraints. The condition (a) provides the basic condition for variables (atomic formulas), and the condition (b) gives the inductive condition. Both clauses show how the defined property behaves step by step from the simplest cases to more complex ones. The smallest set is defined as the set included in all sets that satisfy given conditions; in the case considered, (a) and (b) are the conditions in question. Instead of 'the smallest set such that' one could say $\mathbf{FOR_{PC}}$ is the set satisfying (a), (b) with adding (c) nothing else is a **PC**-formula except objects satisfying (a) and (b). Note that although $\mathbf{AL_{PC}}$ is infinite, every formula over this alphabet possesses a finite length. The full definition of wffs should also indicate how parentheses are to be employed. For example, to be correct, we should write $\neg(A) \in \mathbf{FOR_{PC}}$ rather, than $\neg A \in \mathbf{FOR_{PC}}$, or $(A \Rightarrow B) \in \mathbf{FOR_{PC}}$, not and $A \Rightarrow B \in \mathbf{FOR_{PC}}$. Another simplification is to drop quotes and writing $A \in \mathbf{FOR_{PC}}$ instead of 'A' $\in \mathbf{FOR_{PC}}$. As a matter of fact, I should use quotes, that is, write '$\neg(A)$' $\in \mathbf{FOR_{PC}}$ instead $\neg(A) \in \mathbf{FOR_{PC}}$. The letters A and B do not belong to $\mathbf{L_{PC}}$. The expressions p_1, $p_2 \Rightarrow p_{100}$, $\neg p_4 \Rightarrow (p_4 \wedge p_4)$ are examples of well-formed formulas (the last can be even simplified to $\neg p_4 \Rightarrow p_4 \wedge p_4$, but the inscriptions $p_2 \Rightarrow$, $\neg$, $\Rightarrow p_4$, $\wedge\, p_4$ are not correctly formed and, thereby, entirely meaningless. The letters A and B (possibly also additional capitals (it would be more correct to use A with subscripts) are metavariables or schematic variables. We can think about the letter A as representing an arbitrary wff of **PC**. The expression $A \Rightarrow B$ represents all implications in to $\mathbf{L_{PC}}$, for example—the formulas $p_2 \Rightarrow p_{100}$, $p_4 \Rightarrow p_4 \wedge p_4$, $(p_2 \Rightarrow p_{100}) \Rightarrow (\neg p_4 \Rightarrow p_4 \wedge p_4)$. The subscripts indicate that $\mathbf{AL_{PC}}$ has a denumerably infinite set of variables. A customary manner of representing **PC** employs the letters p, q, r, … as propositional variables. Theoretically speaking, that is incorrect, because every human alphabet consists of a finite number of signs. I use, in accordance with tradition, 'propositional' although 'sentential' would more proper for the already indicated choice of truth-bearers as sentences.

If we say that the expression $A \Rightarrow B$ is a general metalogical representation of all implications constructible in $\mathbf{L_{PC}}$, this also means that this formula belongs to metalogic. On the other hand, the formulas $p_2 \Rightarrow p_{100}$ (or $p \Rightarrow q$) is expressed in the language of logic. Also (**Df1**) and other similar conventions belong to metalogic. Note that logical constants have the same symbols in logic and metalogic. It is possible to codify logic by metalogical schemata instead of employing concrete formulas, and this way of introducing logic will be used in the sequel. This is also called the Hilbert-style formalization of logical calculi which does not need propositional variables with exception of defining $\mathbf{L_{PC}}$. This way of formalizing logic enables us to dispense with the rule of substitution for propositional variables as well as with simplifying the use of quotes (see Chap. 3, Sect. 3.8).►

Now I proceed to presenting **PC** as a formal axiomatic theory,

Axiom schemata:

(**PCA1**) $A \Rightarrow (B \Rightarrow A)$;
(**PCA2**) $(A \Rightarrow (A \Rightarrow B)) \Rightarrow (A \Rightarrow B)$;
(**PCA3**) $(A \Rightarrow B) \Rightarrow ((B \Rightarrow C) \Rightarrow (A \Rightarrow C)$;
(**PCA4**) $A \wedge B \Rightarrow B$;
(**PCA5**) $A \wedge B \Rightarrow B$;

(**PCA6**) $(A \Rightarrow B) \Rightarrow ((A \Rightarrow C) \Rightarrow (A \Rightarrow B \wedge C))$;
(**PCA7**) $A \Rightarrow A \vee B$;
(**PCA8**) $B \Rightarrow A \vee B$;
(**PCA9**) $(A \Rightarrow C) \Rightarrow ((B \Rightarrow C) \Rightarrow (A \vee B \Rightarrow C))$;
(**PCA10**) $(A \Leftrightarrow B) \Rightarrow (A \Rightarrow B)$;
(**PCA11**) $(A \Leftrightarrow B) \Rightarrow (B \Rightarrow A)$;
(**PCA12**) $(A \Rightarrow B) \Rightarrow ((B \Rightarrow A) \Rightarrow (A \Leftrightarrow B)$;
(**PCA13**) $(A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)$;
(**PCA14**) $A \Rightarrow \neg\neg A$;
(**PCA15**) $\neg\neg A \Rightarrow A$;

The rules of inference: (**MP**; modus ponens)

$$\frac{(A \Rightarrow B),\ A}{B}$$

(**DG4**) The above set of axioms has a very nice feature. We can stratify (**PC1**)–(**PC15**) into subsets related to the properties of particular connectives that occur in **PC**. In particular, (**PCA1**)–(**PCA3**) characterize implication, (**PCA4**)–(**PCA6**) regulates the behaviour of conjunction, (**PCA7**)–(**PCA9**) show how disjunction behaves, (**PCA10**)–(**PCA12**) define equivalence, and, (**PCA13**)–(**PCA15**) explain negation. These subsets are useful for extracting some weaker logics being parts of **PC**. For example, if we drop axiom (**PCA15**), we obtain intuitionistic logic (note, however, that, according to the intuitionists, one should not say that intuitionistic logic is a proper part of the classical system). Propositional calculus can be also characterized by many other sets of axiom-schemata. For example, the set $\{(A \Rightarrow B) \Rightarrow ((B \Rightarrow C) \Rightarrow (A \Rightarrow C)), A \Rightarrow (\neg A \Rightarrow B), (\neg A \Rightarrow A) \Rightarrow A\}$ forms the axiomatic base for **PC** due to Łukasiewicz. The first axiom shows that implication is transitive, the second—that contradictions entail everything, and the third—that implication with a false antecedent or true consequent is true. These axioms assume that implication and negation define all other propositional connectives (the same holds for negation and conjunction; there are also other cases as well). For instance, $A \wedge B$ is defined by $\neg(A \Rightarrow \neg B)$, $A \vee B$ by $\neg A \Rightarrow B$, and $A \Leftrightarrow B$ by $(A \Rightarrow B) \wedge (B \Rightarrow A)$. There are also single connectives that define all the others (so-called Sheffer's strokes, see (6) below). If someone thinks that a logical system should be presented in the highest possible economy of expressive means, Sheffer's strokes are important, but I will not enter into details.►

Logic as a theory (*logica docens*) is a collection of theorems. Logical theorems include, firstly, axioms (for instance, (**PC1**)–(**PC15**) and, secondly, formulas provable from axioms by using (**MP**) and formerly proved items. Accordingly, the concept of proof inside **PC** is captured by

(**Df2**) A finite sequence $\boldsymbol{\Sigma}$ of formulas $A_1, \ldots, A_n$ is a proof of a formula A in **PC** (in symbols: $A_1, \ldots, A_n \vdash_{\mathbf{PC}} A$) if and only if
(a) $A = A_n$;

(b) for any A_k ($1 \leq k \leq n$), $A_k \in$ {(**PCA1**)–(**PCA15**)} or $\mathbf{\Sigma}$ contains formulas A_i, A_j ($i, j < k$) such that $A_j = (A_i \Rightarrow A_k)$, provided that $\vdash_{\mathbf{PC}} A_i$ and $\vdash_{\mathbf{PC}} A_j$.

Condition (b) says that A_k is provable from A_i, A_j by the rule (**MP**). Moreover, every element of Σ is provable as either one of the axioms or an already proved theorem. However, if logic is conceived as an instrument of proving theorems in an arbitrary domain of knowledge, that is, if it functions as *logica utens*, a more general concept of proof is needed. Let **X** be an arbitrary set of formulas. The next definition establishes, what it means for a formula A to be provable (by **PC**) from a set **X** of formulas taken as assumptions, independently of whether they belong to logic or not:

(**Df3**) $\mathbf{X} \vdash_{\mathbf{PC}} A$ if and only if
(a) $A \in \mathbf{PC}$ (that is, A is a theorem of **PC**); or
(b) $A \in \mathbf{X}$; or
(c) there is a finite sequence $\mathbf{\Sigma} = A_1, \ldots, A_n$ such that $A = A_n$ and $\mathbf{\Sigma}$ contains formulas
A_i, A_j ($i, j < k$) such that $A_j = (A_i \Rightarrow A_k)$, provided that $\mathbf{X} \vdash_{\mathbf{PC}} A_i$ and $\mathbf{X} \vdash_{\mathbf{PC}} A_j$.

The formulas in **X** do not need to be logical theorems. If $\mathbf{X} = \{A, A \Rightarrow B\}$, we immediately obtain that $\mathbf{X} \vdash_{\mathbf{PC}} B$, but without assuming that the formulas A and $A \Rightarrow B$ are theorems of **PC**. On the other hand, the sequence $\mathbf{\Sigma} = \{A \Rightarrow (A \Rightarrow A)$, $(A \Rightarrow (A \Rightarrow A)) \Rightarrow (A \Rightarrow A)$, $A \Rightarrow A\}$ constitutes a proof of the last formula from the preceding ones. The first formula in $\mathbf{\Sigma}$ is an instance of axiom (**PCA1**), and the second falls under the axiom (**PCA2**). The double application of (**MP**) gives the third formula. Observe also that **X** does not need to be a finite set, although, according to (**Df2**) and (**Df3**), proofs are always finite in their length. It is not forced by logic itself, because there is no problem with speaking on infinite proofs and formulas, but rather a credit for human inferential possibilities).

The concept of proof established by (**Df2**) and that determined by (**Df3**) (sometimes the term 'derivation' is used in this case) are closely related (more precisely, the former is a special case of the latter). This is shown by the deduction theorem (one of the more important theorems in metalogic; roughly speaking, it establishes a parallelism between theorems and rules of inference):

(**DT**) $\{A_1, \ldots, A_n\}_{\subseteq \mathbf{X}} \vdash_{\mathbf{PC}} A$ if and only if $\vdash_{\mathbf{PC}} A_1 \Rightarrow (\ldots \Rightarrow (A_n \Rightarrow A)\ldots)$.

The notation $\{A_1, \ldots, A_n\}_{\subseteq \mathbf{X}}$ indicates that the formulas occurring in the sequence $<A_1, \ldots, A_n>$ belong to the set **X** (the subscript $_{\subseteq \mathbf{X}}$ serves here as a parameter). In fact, (**DT**) can be stated without any relativisation to the set **X**, because every sequence of the type $\mathbf{\Sigma}$ belongs to a set of sentences (in **PC**, we can use the terms 'formula' and 'sentence' synonymously). However, since (**Df3**) explicitly mentions the set **X**, this fact is also displayed in (**DT**). Roughly speaking, the structure of proofs is constituted by theorems of logic. In particular, every logical theorem can be transformed into a correlated rule of inference. For example, the axiom (**PCA4**) justifies the rule $\{A, B\} \vdash_{\mathbf{PC}} A$ (we can also write: $A, B \vdash_{\mathbf{PC}} A$).

This way of looking at logic leads to so-called systems of natural deduction consisting entirely of rules of inference, but without axioms and theorems. Although both concepts of proof are related, their distinction, as we will see, leads to two pictures of logic, which—proves helpful in analysing some problems in the philosophy of logic. Theoretically speaking, logic could be conceived as a body consisting only of theorems, without any rule of inference. Clearly, logic in such a codification would have no practical significance in conducting proofs. This remark shows the real significance of rules of inference in the functioning of logical systems as manuals of tools for deductive inference.

B. **Propositional Calculus: Semantics**

We assume (**BI**) that is the principle of bivalence. Thus, we our set of logical values is $\{\mathbf{1}, \mathbf{0}\}$, where **1** refers to truth and **0** to falsehood. Firstly, we define the valuation function $\mathbf{v^{PC}}$ that acts on propositional variables by

(1) $\mathbf{v^{PC}}$: $\mathbf{Var^{P}} \longrightarrow \{\mathbf{1}, \mathbf{0}\}$.

This means (I will omit superscripts in $\mathbf{v^{PC}}$):

(2) For any $p_i \in \mathbf{VAR^{PC}}$, $\mathbf{v}(p_i) = \mathbf{1}$ or $\mathbf{v}(p_i) = \mathbf{0}$.

According to (1) and (2), the function **v** ascribes logical values, that is, truth or falsehood to every propositional variable. Since the set of variables is denumerably infinite, we have uncountably many valuations, although only a finitely many of their elements are of importance, due to the finite length of formulas. Here are two arbitrary examples of how the valuation function $\mathbf{v_i}$ acts on the first few variables (ordered by their indexes): $\mathbf{v_i}\ (p_1) = \mathbf{1}$, $\mathbf{v_i}\ (p_2) = \mathbf{0}$, $\mathbf{v_i}\ (p_3) = \mathbf{0}$, $\mathbf{v_i}\ (p_4) = \mathbf{1}$, $\mathbf{v_i}\ (p_5) = \mathbf{1}$, …, and $\mathbf{v_j}\ (p_1) = \mathbf{0}$, $\mathbf{v_j}\ (p_2) = \mathbf{1}$, $\mathbf{v_j}\ (p_3) = \mathbf{1}$, $\mathbf{v_j}\ (p_4) = \mathbf{0}$, $\mathbf{v_j}\ (p_5) = \mathbf{1}$, …

The function **v** can be extended to the valuation function to the mapping:

(3) $\mathbf{V^{PC}}$: $\mathbf{FOR^{PC}} \longrightarrow \{\mathbf{1}, \mathbf{0}\}$.

which maps the set $\mathbf{FOR^{PC}}$ onto the set of logical values $\{\mathbf{1}, \mathbf{0}\}$. The full definition **V** for **PC** proceeds as follows (I simplify the notation; 'iff' (the abbreviation for 'if and only if') replaces ⇔ in the last two lines—the same convention is applied in some other places where the symbol ⇔ occurs):

(Df4) (variables) if $A = p_i$, then $\mathbf{V} = \mathbf{v}$;

(¬) if $A = (\neg B)$, then $\mathbf{V}(A) = \mathbf{1} \Leftrightarrow \mathbf{V}(B) = \mathbf{0}$; if $A = (\neg B)$, then $\mathbf{V}(A) = \mathbf{0} \Leftrightarrow \mathbf{V}(B) = \mathbf{1}$;

(∧) if $A = (B \wedge C)$, then $\mathbf{V}(A) = \mathbf{1}$ iff $\mathbf{V}(B) = \mathbf{V}(C) = \mathbf{1}$;
if $A = (B \wedge C)$, then $\mathbf{V}(A) = \mathbf{0} \Leftrightarrow \mathbf{V}(B) = \mathbf{0}$ or $\mathbf{V}(C) = \mathbf{0}$;

(∨) if $A = (B \vee C)$, then $\mathbf{V}(A) = \mathbf{1} \Leftrightarrow \mathbf{V}(B) = \mathbf{1}$ or $\mathbf{V}(C) = \mathbf{1}$; if $A = (B \vee C)$, then
$\mathbf{V}(A) = \mathbf{0} \Leftrightarrow \mathbf{V}(B) = \mathbf{V}(C) = \mathbf{0}$;

(⇒) if $A = (B \Rightarrow C)$, then $\mathbf{V}(A) = \mathbf{1} \Leftrightarrow \mathbf{V}(B) = \mathbf{0}$ or $\mathbf{V}(C) = \mathbf{1}$; if $A = (B \Rightarrow C)$, then
$\mathbf{V}(A) = \mathbf{0} \Leftrightarrow \mathbf{V}(B) = \mathbf{1}$ and $\mathbf{V}(C) = \mathbf{0}$;

($\Leftrightarrow$) if $A = (B \Leftrightarrow C)$, then $\mathbf{V}(A) = \mathbf{1}$ iff $\mathbf{V}(B) = \mathbf{V}(C)$; if $A = (B \Leftrightarrow C)$, then $\mathbf{V}(A) = \mathbf{0}$
iff $\mathbf{V}(B) \neq \mathbf{V}(C)$.

In fact, the last definition reproduces the content of the well-known truth-tables used for checking of validity of **PC**-formulas. The particular lines of (**Df4**) say: (variables) every atomic **PC**-formula is true or false; ($\neg$) negation of a formula is true if and only if the negated formula is false—otherwise, this negation is false; ($\wedge$) conjunction of two formulas is true if and only if both formulas are true—otherwise, the conjunction is false; ($\vee$) disjunction of two formulas is true if and only if at least one of the disjuncts is true—otherwise, the disjunction is false; ($\Rightarrow$) implication of two formulas is true if and only if its antecedent is false or its consequent is true—otherwise, the implication is false; ($\Leftrightarrow$) equivalence of two formulas is true if and only if its components have the same logical value; otherwise the equivalence is false. The inductive character of (**Df4**) and the fact that **v** and **V** are functions in the precise mathematical sense are crucial, because these very circumstances decide that the values of compound formulas depend of valuations of their subformulas. Hence, all linguistic contexts admissible in **PC** as separate units are extensional. We can also express this property by pointing out that the semantics of **PC** is compositional. Moreover, we have here a strict parallelism between compositionality of semantics and compositionality of syntax, because values of expressions in **PC** are calculated according to the structure of its well-formed formulas.

(**DG5**) Some remarks about the problem of compositionality are in order here. Syntactic and semantic compositionality (both described as above) does not exhaust all possibilities. One can also speak about compositionality of meanings understood intensionally (as Fregean senses, Carnap's intensions, uses of expressions, propositions as meaning of sentences, etc.). Compositionality as understood in the context of **PC** (and predicate logic too; see below) concerns syntax and semantics as the theory of reference (see Chap. 6, Sect. 6.2 for general remarks about semantics). Hence, I adopt the view that references are compositional. On the other hand, it is a highly controversial question whether meanings obey the principle of compositionality. Arguments that it is the case seem to be based on an assumption that meanings behave functionally, that is, are functions (in the mathematical sense) of its constituents. If this assumption is given up, the universal validity of compositionality can be soundly criticized. Firstly, natural language seems to lack compositionality. Secondly, even if we limit our interests to languages studied in formal logic, the principle of compositionality excludes intensional contexts from exact mathematical analysis. Both challenges are serious. I do not deny that there is a clear temptation to reject the requirement of compositionality and thereby extend the field of formal semantics.

Yet something can be said in defence of compositional syntax and semantics at least in the context of many languages currently used in various systems of formal logic. Clearly, difficulties with handling intensional contexts by exact mathematical tools can be explained by the lack of compositionality as an essential property of languages. Hence, it may be the case that the principle of compositionality marks a non-trivial border of application of some very important methods—for example, inductive definitions and inductive proofs—that seem indispensable in the formal treatment of any linguistic data and should be applied in handling syntactic as well as semantic features of languages. As far as natural languages are concerned, it seems that only their idealizations become suitable for exact mathematical analysis. This suggests that real natural language cannot be treated by the same methods as logical languages. Thus, we should rather try to extensionalize formal semantics than reject the principle of compositionality. The above (and simplified) remarks are not intended as a general protest against non-compositional syntax or semantics. All attempts to extend mathematical methods in order to develop a semantical theory of intensional contexts (see Parsons 2016 for a survey) or a semantics for independent-friendly logic (see Hintikka 1996, Hintikka 1998, Mann, Sandu, Sevenster 2011) should be welcomed as important proposals. Consequently, there is no reason for unconditional limiting formal semantics to the model theory of classical logic. However, I think that compositionality is too often uncritically complained. To be clear, this defence of compositionally is restricted to special cases.►

Another way of fixing semantics for **PC** consists in considering mappings from $\{\mathbf{1}, \mathbf{0}\}$ into itself. For such mappings, we have the formulas:

(4)

(a) $\{\mathbf{1}, \mathbf{0}\} \longrightarrow \{\mathbf{1}, \mathbf{1}\}$;
(b) $\{\mathbf{1}, \mathbf{0}\} \longrightarrow \{\mathbf{1}, \mathbf{0}\}$;
(c) $\{\mathbf{1}, \mathbf{0}\} \longrightarrow \{\mathbf{0}, \mathbf{1}\}$;
(d) $\{\mathbf{1}, \mathbf{0}\} \longrightarrow \{\mathbf{0}, \mathbf{0}\}$.

Clearly, (c) is similar to negation, but it is not a mapping from formulas onto logical values, but from logical values to logical values. Hence, one must distinguish between negation and this new entity determined by (c). Denote it by the symbol $\neg^*$. Consider now all possible mappings from non-empty subsets of $\{\mathbf{1}, \mathbf{0}\} \times \{\mathbf{1}, \mathbf{0}\}$ onto non-empty subsets of $\{\mathbf{1}, \mathbf{0}\}$, that is, from subsets of the Cartesian product of the set of logical values by itself onto the subsets of the set of logical values. Formally speaking, it is the mapping

(5) $\{\{\mathbf{1},\mathbf{0}\} \times \{\mathbf{1},\mathbf{0}\}\}^2 \rightarrow \{\mathbf{1},\mathbf{0}\}^2$.

This generates 16 combinations (to complete one of earlier remarks, (6e) and (o) represent Sheffer's strokes:

(6) (a) $\{\mathbf{1}, \mathbf{1}\} \longrightarrow \{\mathbf{1}\}, \{\mathbf{0}, \mathbf{1}\} \longrightarrow \{\mathbf{1}\}, \{\mathbf{1}, \mathbf{0}\} \longrightarrow \{\mathbf{1}\}, \{\mathbf{0}, \mathbf{0}\} \longrightarrow \{\mathbf{1}\}$;

(b) $\{\mathbf{1}, \mathbf{1}\} \longrightarrow \{\mathbf{1}\}, \{\mathbf{0}, \mathbf{1}\} \longrightarrow \{\mathbf{1}\}, \{\mathbf{1}, \mathbf{0}\} \longrightarrow \{\mathbf{1}\}, \{\mathbf{0}, \mathbf{0}\} \longrightarrow \{\mathbf{0}\}$;

(c) $\{\mathbf{1}, \mathbf{1}\} \longrightarrow \{\mathbf{1}\}, \{\mathbf{0}, \mathbf{1}\} \longrightarrow \{\mathbf{1}\}, \{\mathbf{1}, \mathbf{0}\} \longrightarrow \{\mathbf{0}\}, \{\mathbf{0}, \mathbf{0}\} \longrightarrow \{\mathbf{1}\}$;

(d) $\{\mathbf{1}, \mathbf{1}\} \longrightarrow \{\mathbf{1}\}, \{\mathbf{0}, \mathbf{1}\} \longrightarrow \{\mathbf{0}\}, \{\mathbf{1}, \mathbf{0}\} \longrightarrow \{\mathbf{1}\}, \{\mathbf{0}, \mathbf{0}\} \longrightarrow \{\mathbf{1}\}$;

(e) $\{\mathbf{1}, \mathbf{1}\} \longrightarrow \{\mathbf{0}\}, \{\mathbf{0}, \mathbf{1}\} \longrightarrow \{\mathbf{1}\}, \{\mathbf{1}, \mathbf{0}\} \longrightarrow \{\mathbf{1}\}, \{\mathbf{0}, \mathbf{0}\} \longrightarrow \{\mathbf{1}\}$;

(f) **{1, 1}** $\longrightarrow$ **{1}**, **{0, 1}** $\longrightarrow$ **{1}**, **{1, 0}** $\longrightarrow$ **{0}**, **{0, 0}** $\longrightarrow$ **{0}**;

(g) **{1, 1}** $\longrightarrow$ **{1}**, **{0, 1}** $\longrightarrow$ **{0}**, **{1, 0}** $\longrightarrow$ **{0}**, **{0, 0}** $\longrightarrow$ **{1}**;

(h) **{1, 1}** $\longrightarrow$ **{1}**, **{0, 1}** $\longrightarrow$ **{0}**, **{1, 0}** $\longrightarrow$ **{1}**, **{0, 0}** $\longrightarrow$ **{0}**;

(i) **{1, 1}** $\longrightarrow$ **{0}**, **{0, 1}** $\longrightarrow$ **{0}**, **{1, 0}** $\longrightarrow$ **{1}**, **{0, 0}** $\longrightarrow$ **{1}**;

(j) **{1, 1}** $\longrightarrow$ **{1}**, **{0, 1}** $\longrightarrow$ **{1}**, **{1, 0}** $\longrightarrow$ **{0}**, **{0, 0}** $\longrightarrow$ **{0}**;

(k) **{1, 1}** $\longrightarrow$ **{0}**, **{0, 1}** $\longrightarrow$ **{1}**, **{1, 0}** $\longrightarrow$ **{1}**, **{0, 0}** $\longrightarrow$ **{0}**;

(l) **{1, 1}** $\longrightarrow$ **{1}**, **{0, 1}** $\longrightarrow$ **{0}**, **{1, 0}** $\longrightarrow$ **{0}**, **{0, 0}** $\longrightarrow$ **{0}**;

(m) **{1, 1}** $\longrightarrow$ **{0}**, **{0, 1}** $\longrightarrow$ **{1}**, **{1, 0}** $\longrightarrow$ **{0}**, **{0, 0}** $\longrightarrow$ **{0}**;

(n) **{1, 1}** $\longrightarrow$ **{0}**, **{0, 1}** $\longrightarrow$ **{0}**, **{1, 0}** $\longrightarrow$ **{1}**, **{0, 0}** $\longrightarrow$ **{0}**;

(o) **{1, 1}** $\longrightarrow$ **{0}**, **{0, 1}** $\longrightarrow$ **{0}**, **{1, 0}** $\longrightarrow$ **{0}**, **{0, 0}** $\longrightarrow$ **{1}**;

(p) **{1, 1}** $\longrightarrow$ **{0}**, **{0, 1}** $\longrightarrow$ **{0}**, **{1, 0}** $\longrightarrow$ **{0}**, **{0, 0}** $\longrightarrow$ **{0}**.

Now, symbolic denominations $\wedge$*, $\vee$*, $\Rightarrow$*, $\Leftrightarrow$* for (h), (e), (f) and (j) respectively are justified, because (h) corresponds to conjunction, (b) to disjunction, (c) to implication, and (g) to equivalence. Returning to (**Df4**), (e) and (o) are the so-called Sheffer functions, that is, connectives sufficient to express the rest of the logical constants of **PC**.

The matrix for **PC** is the entity **MAT$^{\mathbf{PC}}$ = <LV = {1, 0}**, **DV = {1}**, ¬*, $\wedge$*, $\vee$*, $\Rightarrow$*, $\Leftrightarrow$***}**, where **{1**, **0}** is the set of logical values, **{1}** is the set of designated values, and

(7) (a) ¬* is a unary function from **LV** to **LV** such that ¬*(**1**) = **0** and ¬*(**1**) = **0**;
(b) $\wedge$* is a binary function from **LV** $\times$ **LV** to **LV** such that $\wedge$*(**1**, **1**) = **1**, $\wedge$* (**0**, **1**) = **0**,
$\wedge$*(**1**, **0**) = **0**, $\wedge$*(**0**, **0**) = **0**;
(c) $\vee$* is a binary function from **LV** $\times$ **LV** to **LV** such that $\vee$*(**1**, **1**) = **1**, $\vee$* (**1**, **1**) = **1**,
$\vee$*(**1**, **0**) = **1**, $\vee$* (**0**, **0**) = **1**;
(d) $\Rightarrow$* is a binary function from **LV** $\times$ **LV** to **LV** such that $\Rightarrow$*(**1**, **1**) = **1**, $\Rightarrow$*(**0**, **1**) = **1**, $\Rightarrow$*(**1**, **0**) = **0**, $\Rightarrow$*(**0**, **0**) = **1**;
(e) $\Leftrightarrow$* is a binary function from **LV** $\times$ **LV** to **LV** such that $\Leftrightarrow$*(**1**, **1**) = **1**,

MAT$^{\mathbf{PC}}$ is then an algebra of logical values. **PC** is about **MAT$^{\mathbf{PC}}$** in the sense that its formulas always have values from **LV**. Further, a formula *A* is valid in **MAT$^{\mathbf{PC}}$** if and only if its value always belongs to **DV**. The algebra **MAT$^{\mathbf{PC}}$** can be considered as the world of logic, that is, the reality which logic is about. It is a very poor world, because it consists of two objects and twenty binary relations.

Having introduced valuation (I will use this apparatus), we can define several semantic concepts for **PC**. The most crucial are collected in

(**Df5**) (**a**) A is PC-satisfiable (has a PC–model) ⇔ for some **V**, **V**(A) = **1**;
(**b**) **X** is PC-satisfiable ⇔ for every $A \in$ X, A is PC–satisfiable;
(**c**) A is a PC-tautology ($A \in \mathrm{TL}_{\mathrm{PC}}$, $\models_{\mathrm{PC}} A$) ⇔ for every **V**, **V**(A) = **1**;
(**d**) A is a **PC**-semantic consequence of **X** (**X** $\models_{\mathbf{PC}} A$) ⇔ for every **V**, if **X** is **PC**-satisfied by **V**, so is A.
(**e**) A is a **PC**-contradiction ($A \in \mathrm{CTL}_{\mathrm{PC}} \mathrel{=\!\!|}_{\mathbf{PC}} A$) ⇔ A is not **PC**–satisfiable, that is, for any **V**, **V**(A) = **0**.

According to (**Df5c**), we can say that tautologies are true under all possible valuations, or true in all possible worlds (every **v** can be considered as a representation of a possible world), or true in all state-descriptions, or true in all circumstances. The concept of tautology corresponds to that of validity in $\mathbf{MAT}^{\mathbf{PC}}$.

(**DG6**) The concepts of syntax and semantics still wait for an explanation in this book. The matter will be discussed in Chap. 6, Sects. 6.2–6.3. I took this route quite deliberately. At first, I wanted to give examples of syntactic and semantic considerations that occur in pure logic. Roughly speaking, the concepts of formula, axiom (axiom-scheme), rule of inference, proof and theorem, are syntactic, but the concepts of valuation and tautology are semantic. A comparison of both kinds of ideas is the basic task of metalogic.►

(**DG7**) (**Df4**) has a certain troublesome feature, particularly for philosophers. It is clear that explanations of truth-conditions for particular kinds of compound sentences (negations, conjunctions, disjunctions, implications and equivalence, in order to limit thus remarks to the given axioms of **PC**)—implicitly or even explicitly—use linguistic counterparts of propositional connectives in ordinary language. Consider, for example, the condition (∨). It says that disjunction of two sentences A and B is true if A is true or B is true. However, this condition explains the meaning of 'or' in the prescribed way dictated by the tasks of logic. It is difficult to find a fully satisfactory answer, acceptable to those philosophers, who, like the phenomenologists, claim that any philosophical should start as presupposition-free. Let me only note that, strictly speaking, 'or' and other connectives (more precisely, the words functioning as their ordinary counterparts) used in (**Df4**) do not belong to $\mathbf{L}_{\mathbf{PC}}$. Hence, one can say that we use them in a minimal intuitive meaning, sufficient for stating the truth-conditions for the connective ∨ as it is employed in propositional calculus. Anyway, I will use the same symbols in metalogic and logic, even though in many cases ordinary words occur in metalogical formulations (for example, 'or' for the disjunction).►

C. First-Order Predicate Calculus: Syntax

The alphabet of $\mathbf{AL}_{\mathbf{FOL}} = \mathbf{VAR}^{\mathbf{FOL}} \cup \mathbf{CO}^{\mathbf{FOL}} \cup \mathbf{FUN}^{\mathbf{FOL}} \cup \mathbf{CONST}^{\mathbf{FOL}} \cup \mathbf{PRED}^{\mathbf{FOL}} \cup \{(,)\}$, where $\mathbf{VAR}^{\mathbf{FOL}}$—is a possibly denumerable infinite set of individual variables $x_1, x_2, x_3, \ldots$; $\mathbf{CO}^{\mathbf{FOL}}$—a **c**ountable (denumerably infinite, finite or empty set) of individual constants $a_1, a_2, a_3, \ldots$; $\mathbf{FUN}^{\mathbf{FOL}}$—a countable (denumerably infinite, finite or empty) set of function letters $f_{1(1)}, f_{1(2)}, \ldots. f_{2(1)}, f_{2(2)},$

…, $f_{i(j)}$, … (in the symbol $f_{i(j)}$, the index i indicates the arity of a function letter and the index j its place in the list of function letters of a given arity; for example, the symbol $f_{1(2)}$ refer to the second function letter of arity 1); $\mathbf{CONST^{FOL}}$–$\mathbf{CONST^{PC}}$ ∪ {∀ (universal quantifier—'for any', 'for every', for all'), ∃ (existential quantifier —'for some', 'there is', 'there exists'); = (identity; 'is identical with', 'is the same as')}; $\mathbf{PRE^{FOL}}$—a possibly denumerable infinite set of predicate letters $P_{1(1)}$, $P_{1(2)}$, …, $P_{2(1)}$, $P_{2(2)}$, …, $P_{i(j)}$, … (in the symbol $P_{i(j)}$, the index i indicates the arity of a predicate letters (briefly: a predicate), and the index j—its place in the list of predicates of a given arity (for example the symbol $P_{1(2)}$ refers to the second predicate of arity 1); {(,)}—as in the case of **PC**. Moreover, $\mathbf{TER^{FOL}}$ (the set of terms of **FOL**) = $\mathbf{VAR^{FOL}} \cup \mathbf{CO^{FOL}} \cup \mathbf{FUN^{FOL}}$.

(**Df6**) The set $\mathbf{FOR^{FOL}}$ of formulas constructible over $\mathbf{AL_{FOL}}$ is the smallest set satisfying if $t_1, \ldots, t_k \in \mathbf{TER^{FOR}}$, then $Pj(t_{1(j)}, \ldots, t_{k(j)}) \in \mathbf{FOR^{FOL}}$;
(**a**) if $t_i, t_k \in \mathbf{TER^{FOR}}$, then $(t_i = t_k) \in \mathbf{FOR^{FOL}}$;
(**b**) if $A, B \in \mathbf{FOR^{FOL}}$, and $v_i \in \mathbf{VAR^{FOR}}$, then $\forall v_i(A(v_i) \in \mathbf{FOR^{FOL}}$, $\exists v_i\,(A(v_i) \in \mathbf{FOR^{FOL}}$,$\neg A \in \mathbf{FOR^{FOL}}$, $A \wedge B \in \mathbf{FOR^{FOL}}$, $A \vee B \in \mathbf{FOR^{FOL}}$, $A \Rightarrow B \in \mathbf{FOR^{FOL}}$, $A \Leftrightarrow B \in \mathbf{FOR^{FOL}}$.

(**Df7**) (**a**) If $A = \forall v_i\, B$ or $A = \exists v_i B$, then the formula B is the scope of ∀ or ∃;
(**b**) if $v_i = v_j$ and $A = \forall v_i B(v_j)$ or $A = \exists v_i B(v_j)$, then the occurrence of variable v_j is bound by ∀ or ∃ located in front of B; otherwise, the occurrence of v_j is free;
(**c**) if the variable v_i does not occur in A, then $\forall v_i(A(v_i) = A)$ and $\exists v_i\,(A(v_i) = A)$;
(**d**) a sentence is a formula without the occurrence of free variables; otherwise, it is an open formula (the notation $A \in \mathbf{SEN}$ means 'A is a sentence');
(**e**) the term t (the letter t is a metavariable for terms) is substitutable for a variable v in a formula A if and only v has free occurrence in A, and no free occurrence of a variable in A before substitution is bound after substitution ('$A(t/v)$' means that 't is substituted for v in A').

Axiom-schemata:

(**FOLA0**) All **PC**-axioms, provided that schematic letters represent **FOL**-formulas;
(**FOLA1**) $\forall v(A(v)) \Rightarrow A(t/v)$, provided that t is substitutable for v in A;
(**FOLA2**) $A(t) \Rightarrow \exists v(A(v))$.
(**FOLA3**) $\forall t_i(t_i = t_i)$;
(**FOLA4**) $(\ldots(s_1 = t_1) \wedge \ldots \wedge (s_n = t_n)\ldots) \Rightarrow (P_{n(i)}(s_{1(i)}, \ldots, s_{n(i)}) \Rightarrow P_{n(i)}(t_{1(i)}, \ldots, t_{n(i)}))$.

Rules of inference:

(**MP**), provided that the schematic letters represent **FOL**-formulas;
(**Ad∀**) (introduction ∀)

$\frac{A \Rightarrow B(v)}{A \Rightarrow \forall v B(v)}$, provided that v is not free in A;

(**Ad**$\exists$) (introduction $\exists$)

$\frac{A(t/v) \Rightarrow B(v)}{A(t) \Rightarrow \exists v B(v)}$, provided that t is substitutable for v.

Most of the comments made in (**DG1**) apply *mutatis mutandi* to **FOL**. In particular, although $\mathbf{AL_{FOL}}$ is infinite, **FOL**-formulas are finite in length. (**Df6a–b**) defines atomic sentences of **FOL**. The sense of the restrictions imposed on substitution will be explained on the occasion of discussing semantic matters in the next subsection. In general, they prevent inferences from true premises to false conclusions. The inclusion of identity in the stock of logical constants is controversial. Also this issue will be discussed in the next subsection. The above presentation of the syntax of **FOL** does not admit mixed formulas, that is, formulas composed of $\mathbf{FOR^{FOL}}$ and propositional formulas—for example, '$A(x) \Rightarrow p$' (in informal remarks, indexes of variables are omitted)—but this restriction can be dropped with determining any additional difficulty. On the other hand, **PC** can be considered as a proper part of **FOL** (see (**FOLA0**)) (in fact it is).

Formulas described in (**Df6a**) and (**Df6b**) are called atomic. Preceding a formula by the universal quantifier forms its universal closure; preceding a formula by the existential quantifier results in its existential closure. (**Df7c**) concerns the so-called vacuous occurrence of a variable in a formula and says that the universal existential closure of a formula is equal to this formula in the case of vacuous occurrences of variables. Function letters and predicates are parameters, not variables, because quantifiers of **FOL** that bind individual variables only. If we introduce quantifiers that predicate letters (of **FOL**), the second-order logic arises (I consider this logic as an illustration of higher-order logic and its problems). For example, the formula $\forall x \exists P\ (P(x))$ is second-order. This shows the relativity of the concept of a variable, because to be a variable means to be bounded by a quantifier of some kind. The axiom (**FOLA4**) expresses the weak version of Leibniz's law; the strong or full one arises when both implications is replaced by equivalences. Intuitively, terms (but not predicates, contrary to ordinary grammar) play the role of names. This explains why the function letters are included into the collection of terms. The arithmetical expression $x + y$ refer to the operations of adding y to x (we can read it 'the sum of x and y'). Thus, a function symbol is a term having other terms as its arguments; in the language of syntactic categories (see Chap. 6, Sect. 6.3), the symbol + is a name-forming functor (the term 'functor' was introduced by Polish logicians as a general name) of two term as its arguments. Theoretically, function symbols are a special kind of predicates. Every n-ary function can be replaced by a suitable (n + 1)-predicate. For example, the expression $x + y$ has its predicate counterpart in the expression $x + y =$ ('x plus y is equal to'). This example provides an illustration that predicates are sentence-forming functors having terms as their arguments. This feature of **FOL** is very different from the account of predicates in traditional (in particular, Aristotelian logic). I will return to this question below.

Although function symbols are dispensable in **FOL**, their role, mostly in mathematics, decides about their common introduction in the abstract presentation of logic of predicates. If functions occur in a language, individual constants can be interpreted as zero-argument (zeroary) function symbols; this convention looks artificially (and certainly is from the ordinary point of view), but is admissible form the theoretical point of view). These considerations lead to the question of which symbols are necessary for having a first-order language. We can drop individual constants, function-symbols, and eventually identity. Thus, variables, predicates and logical constants (other than identity) are indispensable. According to some terminological proposals (see Church 1956, p. 218), **FOL** without constants, function-symbols, and identity can be regarded as pure quantification (functional) calculus. If constants or functions are added, an applied **FOL** arises (identity calls a decision, see below). This terminology has its justification in the fact that constants and functions letters serve to formulate statements about special objects and their behaviour.

The concepts of proof defined in (**Df2**) and (**Df3**) for **PC** are easily (but not automatically) adapted for **FOL** by

(**Df8**) A finite sequence $\mathbf{\Sigma}$ of formulas $A_1, \ldots, A_n$ is a proof of a formula A in **FOL** ($A_1, \ldots, A_n \vdash_{\mathbf{FOL}} A$) iff
(**a**) $A = A_n$;
(**b**) for any A_k ($1 \leq k \leq n$), $A_k \in$ {(**FOLA0**)–(**FOLA4**)}, or
(**i**) the sequence $\mathbf{\Sigma}$ contains formulas A_i, A_j ($i, j < k$) such that $A_j = (A_i \Rightarrow A_k)$, provided that $\vdash_{\mathbf{FOL}} A_i$ and $\vdash_{\mathbf{FOL}} A_j$;
(**j**) there is a formula A_i ($1 \leq j \leq k - 1$) such that A is obtainable from A_i by applications of (**Ad**∀) or (**Ad**∃), provided that $\vdash_{\mathbf{FOL}} A_i$.

(**Df9**) $\mathbf{X} \vdash_{\mathbf{FOL}} A$ if and only if
(**a**) $A \in$ **FOL** (A is a theorem of **FOL**); or
(**b**) $A \in \mathbf{X}$; or
(**c**) there is a finite sequence $\mathbf{\Sigma}$ of formulas $A_1, \ldots, A_n$ such that $A = A_n$ and $\mathbf{\Sigma}$ contains formulas A_i, A_j ($i, j < k$) such that $A_j = (A_i \Rightarrow A_k)$; or
(**d**) there is a formula $A_i \in \mathbf{X}$ such that A is obtainable from A_i by applications of (**Ad**∀) or (**Ad**∃).

The deduction theorem holds for **FOL** if we exclude open formulas (I omit the explanation of this restriction, related to the interplay of free and bound occurrences of variable). This completes the presentation of the syntax of **FOL**. Recalling that, due to (**FOLA0**), **PC** is contained in (is a subsystem of) **FOL**, propositional calculus and first-order predicate calculus form together the system called elementary logic. In order to explain some terminological matters, note that sometimes **FOL** is identified with a system based on (**FOLA0**)–(**FOLA2**) plus (**MP**), (**Ad**∀), and (**Ad**∃). If (**FOLA3**) and (**FOLA4**) are added, the resulting system is first-order logic with identity ($\mathbf{FOL}_{=}$).

D. First-Order Logic: Semantics

The semantic interpretation of sentences only via their logical values is mostly responsible for the noted purity of the world of **PC**. In particular, the internal structure of atomic sentences in the sense of **PC** plays no logical role in ascribing truth or falsehood to them. Every atom of **PC** constitutes a separate unit, which cannot be analyzed any further by logical devices accessible in propositional logic. We ascribe concrete truth values to sentences by understanding their meaning or by conventional fiat, but this is accomplished by employing entirely extralogical devices. Predicate calculus partly changes this situation, because it takes into account the inner structure of sentences and thereby makes possible to speak about the properties of objects and the relations holding between them (note that I am not saying that grasping the meaning is irrelevant). Consider the sentence (a) 'John is tall'. Although one can say that this sentence says that John has a certain property also on the level of **PC**, namely, that he is tall—this qualification is not expressible by resources of propositional calculus, because the grammar of **PC**-sentences has nothing to do with expressions referring to objects and properties. On the other hand—(a) as falling under the form P (a)—pictures structurally the fact that John is tall. Hence, the rules of **FOL** have a deeper link with the world. In particular, they exhibit logical machinery related to quantifiers and their scopes. This enables us to express generality and particularity or, for instance, the rule that generality entails particularity (*dictum de omni et nullo*) that is, if a property is attributed to all items belonging to a given scope, **X**. it is attributable to some items from **X**, and if a property is not possessed by any item, it cannot be attributed to some items. These principles known from the traditional logic, have an exact wording in **FOL**. This remark throws some light on what is going on in saying that there is some progress in logic.

Historically speaking, predicate calculus as a logical system appeared for the first time in Frege, although not as **FOL**. First-order logic was extracted as a separate system in the 1920s. (see Hilbert, Ackermann 1928). The introduction of quantifiers became one the most notable discoveries in the entire history of logic. Logical analysis of quantifiers was preceded by its use in mathematics in making the foundations of analysis sufficiently precise (Cauchy, Weierstrass). Note that Latin has no means to express quantifiers in the modern sense; the same concerns classic Greek, but this language did not play the role of the universal device of scientific communications (except antiquity). Perhaps the development of predicate logic grew as a result of changing Latin as the *lingua franca* of science into national languages. It is important to see how atomic formulas of **FOL** differ from singular sentences of traditional logic. Apparently they have the same form, namely 'a is P'. According to the traditional view 'is' functions as a copula linking the subject term a and a predicate term P, but the modern view considers the expression 'is P' as whole. As far as the matter categorical features are concerned, 'is' as the copula is a sentence forming operator of two arguments, but 'is P', in the light of **FOL**, is a sentence forming operator of one argument.►

The semantics of **FOL** is referential. This means that semantic relations express connections that hold between language and the world. The word 'the world' should be understood *cum grano salis* here. In fact, we have in mind a domain of objects, properties of objects, and relationships holding between them as defined over this domain. It is a quite separate problem of how domains studied in formal semantics are associated with the real world. We need not to be concerned with this question At the moment; I will return to it in Chap. 9, Sect. 9.6. Our first task consists in defining the concept of interpretation the language of **FOL** ($\mathbf{L_{FOL}}$). This will be done in two stages. We first show how to interpret $\mathbf{AL_{FOL}}$, and we then pass to $\mathbf{L_{FOL}}$ (I will omit subscripts in most cases) conceived as a set of sentences. Just as in the case of **PC**, to give an interpretation of a first-order formula means to define its truth-conditions (or simply, truth) in the case of sentences, or its satisfaction-conditions in the case of open formulas. The main difference between **PC** and **FOL** consists in the treatment of atoms and quantified sentences, because the semantic rules of first-order logic take into account the inner structure of sentences, not only their role in truth/falsehood games governed by $\mathbf{V^{PC}}$. Hence, we must to fix how constituents of sentences behave semantically, before we pass to the second stage. Since truth-conditions for sentences depend on how the concept of truth is defined, the full discussion of this issue is postponed to Chaps. 7 and 8; in particular, in Chap. 7, Sect. 7.6 I shall explain the semantic difference between sentences and open formulas. However, at the end of this section I shall give a simplified truth-definition for a special case.

Since the logical behaviour of propositional connectives has already been established in **PC**, their meaning can be taken for granted. Hence, it remains to work with individual variables, individual constants, function-letters and predicates. We think of variables as representing objects, of individual constants as names of concrete objects (individuals), functions letters as names of functions, unary predicates as referring to properties, n-ary predicates ($n \geq 2$) as referring to relations of a suitable arity, and quantifiers as kinds of functions. The linguistic material of **L** has its representation in the union (I omit parentheses as auxiliary signs) $\mathbf{VAR^{FOL}} \cup \mathbf{CO^{FOL}} \cup \mathbf{FUN^{FOL}} \cup \mathbf{CONST^{FOL}} \cup \mathbf{PRE}$. Its closer description is given by:

$$(*)\mathrm{VAR}^{\mathbf{FoL}} \cup \{a_1, a_2, a_3, \ldots\} \cup \left\{f_{1(1)}, f_{1(2)}, \ldots, f_{2(1)}, f_{2(2)}, \ldots f_{i(j)}, \ldots\right\}$$
$$\cup \left\{P_{I(l)}, P_{I(2)}, \ldots, P_{2(1)}, P_{2(2)}, \ldots P_{i(j)}, \ldots\right\} \cup \{=, \forall, \exists\}.$$

In order to define an interpretation of $\mathbf{AL_{FOL}}$ (as presented by (#)), we must have at our disposal an object which in a sense is similar to this alphabet. This is realized by defining the structure

$$\Re = \ <\mathbf{U}, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{f_{1(1)}}, \mathbf{f_{1(2)}}, \ldots, \mathbf{f_{i(j)}}, \mathbf{P_{1(1)}}, \mathbf{P_{1(2)}}, \ldots, \mathbf{P_{i(j)}}, \ldots>,$$

where (a) **U** is a non-empty (no other condition is imposed on cardinality, that is, U can be finite, denumerably infinite or uncountable) set of objects (the carrier of interpretation, the universe of discourse); (b) for every i (i = 1, 2, …), $\boldsymbol{a_i} \in \mathbf{U}$ (we say that $\mathbf{a_i}$ is a distinctive element of **U**) (c) for every i, j (i, j = 1, 2, …), $\mathbf{f_{i(j)}}$: $\mathbf{U} \times \ldots$ (i-times) $\times\ \mathbf{U} \longrightarrow \mathbf{U}$, that is, $\mathbf{f_{i(j)}}$ is a function which maps the Cartesian product of

U taken *i*-times into **U**, and (d) for every *i, j* ($i, j = 1, 2, \ldots$) $\mathbf{P_{i(j)}} \subseteq \mathbf{U} \times \ldots$ (*i*-times) $\times$ **U**. Intuitively, any object of the type $\mathbf{P_{i(j)}}$ is an *i*-ary relation (other symbol: $\mathbf{R}_{\mathbf{i(j)}}$) defined on **U**; according to earlier explanations concerning $\mathbf{AL_{FOL}}$, the sets defined in (a) and (b)—unlike to the set $\{\mathbf{P_{1(1)}}, \mathbf{P_{1(2)}}, \ldots, \mathbf{P_{i(j)}}, \ldots\}$ of predicates—can be empty. If $i = 1$, $\mathbf{P_{1(j)}} \subseteq \mathbf{U}$, and can be called a property (more strictly: a property extensionally interpreted).

(**DG8**) Although **U** is not empty, we should not exclude the empty property being identified with the empty set ∅. Since ∅ is a subset of any set, every universe allows an empty property. Incidentally, this case is a good illustration of the difference between an extensional and an intensional interpretation of properties, because 'is a round square' and 'is the greatest standard natural number' have different meanings (intensions), but refer to the same set, namely the empty set. Hence, we have many intensional empty properties and only the extensional one. Incidentally speaking, the empty set cannot be considered as representing the Nothing (in the ontological sense, for example, exemplified by Heidegger's deliberations about *Das Nichts nichet.* The empty set is just something, for example, the product of two sets not having common elements.►

The structure $\mathfrak{R}$ is suitable for an interpretation of **AL** if both are similar. We know in advance something about **AL**, for example, that it has (or not) some individual constants and/or predicates of some particular arity. The interpretative structure must be similar to an alphabet in the sense that if **AL** distinguishes some constants and has predicates of an arity *n*, $\mathfrak{R}$ should contain in its inventory some entities, for example distinctive individuals and *n*-ary relations. Details are provided by the content (elements) of the sets $\{a_1, a_2, a, \ldots\}$ and $\{P_{1(1)}, P_{1(2)}, \ldots, P_{i(j)}, \ldots\}$. Technically speaking, this decides about the signature of language **L** (see Manzano 1999, pp. 22–23, 54–55 for a fully rigorous treatment; I only sketch here some rudiments in an informal way). Assume that we have two individual constants a_1 and a_2, one unary predicate $P_{i(1)}$ and one binary predicate $P_{j(2)}$. This language has the signature <0; 0; 1; 2> (it is customary to sign constants by 0, and predicates by numerals expressing their arities). If we get information that this sequence represents the signature of a language, we know that our linguistic framework has two constants, one unary predicate and one binary predicate, although we do not know anything more concrete about the relevant items. However, we also know that a proper structure for interpreting of our language **L** must have at least the signature <0*; 1*; 2*> (starred numbers indicate signatures of interpretations). This means that $\mathfrak{R}^{\mathbf{L}}$ contains at least one distinctive constant, one unary predicate and one binary predicate. Why at least but not the same? The answer is that we do not rule out the situation that two (or more) constants name the same objects and two (or more) predicates refer to the same property or relations.

(**DG9**) Example: Let **L** have {'Socrates', 'Plato', 'Aristocles'} as the set **CO** and {'is a philosopher', 'is the teacher of'} as the set **PRED**. Assume that we do not know anything about denotations of particular expressions. We know only that the signature of **L** is <0, 0, 0; 1; 2>. When looking for an interpretation of **L**, we can consider structures with signatures <0*, 0*, 0*; 1*; 2*>, <0*, 0*; 1*; 2*>, and <0*;

1*; 2*>. If our interpretation is to be historically faithful, the structure should have the signature <0*, 0*; 1*; 2*>, because the constants 'Aristocles' and 'Plato' refer to the same person. Now, if the interpretative structure would have no identity relation, the sentences 'Socrates is a teacher of Plato' and 'Socrates is a teacher of Aristocles' could not be interpreted in it. One can claim that a good language should not have any expression referring to different entities. If this requirement is accepted, then perfect symmetry (one-to-one correspondence) holds between signatures of languages and signatures of interpretations. Think also about empty predicates and the empty property in extension as named by different predicates.►

(**DG10**) The signature of any language **L** strongly depends on its **AL**. For instance, the signature of a given **L** determined by the fact whether **AL** has individual constants or function letters. Assume that we consider the pure **FOL**, that is, a language without constants and function letters. Full generality requires that we have a denumerable list of predicates for any arity. The signature of this language must indicate all possible predicates. Hence, it has the form of the infinite sequence **s** = <1, 1, 1, …; 2, 2, 2, …; 3, 3, 3, …; *n*, *n*, *n*, …>. Every element of every subsequence in **s** represents a separate intensional property. Since we cannot rule out the situation of a perfect (one-to-one correspondence) between intensions and extensions, we have uncountably many properties that are expressible in the pure first-order logic. Although this fact has no special relevance to customary, ordinary or even scientific, particularly mathematical, applications of formal logic, it has a considerable significance for the theory of logic.►

The above considerations suggest that three different parameters are involved in every interpretation, namely **AL**, $\mathfrak{R}$, and a correlation between them. The last element is displayed by a valuation function $\mathbf{V}^{\mathbf{FOL}}$, which ascribes semantic values taken from the structure $\mathfrak{R}$ to expressions taken from **AL**. Since the structure $\mathfrak{R}$ is a fairly complicated ordered set, it is convenient to convert it into a family of sets (the notation **X** ÷ **Y** refers to the difference of the sets **X** and **Y** (**X** minus **Y**—the elements of the set **Y** are eliminated from the set **Y**; the symbol **Id** denotes the relation of identity):

$$\Psi^{\mathbf{L},\mathfrak{R}} = \{\mathrm{U}, \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots\}, \{\mathbf{f}_{\mathbf{1}(\mathbf{1})}, \mathbf{f}_{\mathbf{1}(2)}, \ldots, \mathbf{f}_{\mathbf{i}(\mathbf{j})}, \ldots\}, \{\mathbf{P}_{\mathbf{1}(\mathbf{1})}, \mathbf{P}_{\mathbf{1}(\mathbf{2})}, \ldots, \mathbf{P}_{\mathbf{i}(\mathbf{j})}, \ldots\}\}$$

The notation $\Psi^{\mathbf{AL},\mathfrak{R}}$ indicates that the set in question is related to the alphabet **AL** and the structure $\mathfrak{R}$; we can assume that $\Psi^{\mathbf{AL},\mathfrak{R}}$ has the same signature as $\mathfrak{R}$. Formally speaking, an interpretation $\mathfrak{I}$ is the triple <**AL**, $\Psi^{\mathbf{AL},\mathfrak{R}}$, $\mathbf{V}^{\mathbf{FOL}}$>, where $\mathbf{V}^{\mathbf{FOL}}$: **AL** $\longrightarrow \Psi^{\mathbf{AL},\mathfrak{R}}$. More specifically I omit the superscript referring to **FOL**), we have:

(**Df10**) (**a**) $\mathbf{V}^{\mathbf{VAR}}$: **VAR** $\longrightarrow$ **U**;
(**b**) $\mathbf{V}^{\mathbf{CO}}$: **CO** $\longrightarrow \{\mathbf{a_1}, \mathbf{a_2}, \mathbf{a_3}, \ldots\}$;
(**c**) $\mathbf{V}^{\mathbf{FUN}}$: **FUN** $\longrightarrow \{\mathbf{f}_{\mathbf{1(1)}}, \mathbf{f}_{\mathbf{1(2)}}, \ldots, \mathbf{f}_{\mathbf{2(1)}}, \mathbf{f}_{\mathbf{2(2)}}, \ldots, \mathbf{f}_{\mathbf{i(j)}}\}$;
(**d**) $\mathbf{V}^{\mathbf{PRE}}$: **PRE** $\longrightarrow \{\mathbf{P}_{\mathbf{1(1)}}, \mathbf{P}_{\mathbf{1(2)}}, \ldots, \mathbf{P}_{\mathbf{2(1)}}, \mathbf{P}_{\mathbf{2(2)}}, \ldots, \mathbf{P}_{\mathbf{i(j)}}\}$;
(**e**) $\mathbf{V}^{=}$: = $\longrightarrow$ **Id**, for any $\Psi^{\mathbf{AL},\mathfrak{R}}$;

(**f**) $\mathbf{V}^{\forall}$: $\forall \longrightarrow \mathbf{U}$, for any $\boldsymbol{\Psi}^{\mathbf{AL},\mathfrak{R}}$;
(**g**) $\mathbf{V}^{\exists}$: $\exists \longrightarrow \{\mathbf{U}^2 \div \varnothing\}$, for any $\boldsymbol{\Psi}^{\mathbf{AL},\mathfrak{R}}$.

(**Df10**) assumes the meaning of propositional connectives established by PC and explained in (6). Points (**a**)–(**d**) pertain to extralogical expressions, the rest—to logical constants of **FOL**. Thus, the valuation function maps variables onto **U**, individual constants—onto the chosen elements of the set of individual objects constituting the universe of a given discourse, function-letters—onto functions, and predicates—onto suitable (according to the number of arguments, that is, arities) properties and/or relations. If **AL** has no constants or functions, the mappings defined in (**b**) and (**c**) are empty. The behaviour of **V** in the case of logical constants is independent of concrete domains (it is mark of the universality of purely logical concepts). The function **V** is neither injective nor surjective. That **V** is not injective provides a formal reason for why the signatures of **AL** and $\mathfrak{R}$ can be different.

We are now prepared to define the context $\mathbf{v}([\ldots], \mathfrak{I})$ (to be read: 'the value of an expression [...] under the interpretation $\mathfrak{I}$'). It is done by

(**Df11**) (**a**) $\mathbf{v}(v_i, \mathfrak{I}) = \boldsymbol{u_i}$, where $\boldsymbol{u_i} \in \mathbf{U}$;
(**b**) $\mathbf{v}(a_i, \mathfrak{I}) = \boldsymbol{a_i}$;
(**c**) $\mathbf{v}(f_{i(j)}, \mathfrak{I}) = \mathbf{f_{i(j)}}$;
(**d**) $\mathbf{v}(P_{i(\mathrm{j})}, \mathfrak{I}) = \mathbf{P_{i(j)}}$;
(**e**) $\mathbf{v}(t_i = t_j) = \mathbf{Id}(\mathbf{v}(t_i), \mathbf{v}(t_j))$;
(**f**) $\mathbf{v}(\forall) = \mathbf{U}$ (the whole universe is the value);
(**g**) $\mathbf{v}(\exists) = \mathbf{U}'$, where $\mathbf{U}' \in \{\mathbf{U}^2 \div \varnothing\}$ (any non-empty subset of **U** is a value).

Several comments are in order here. I once again remind the reader of three things: (a) that so far we only gave an account of how elements of **AL** are interpreted; (b) that this procedure does not constitute an interpretation of **L** as a set of sentences, because (**Df11**) does not prescribe how truth and falsehood of sentences are related to semantic values of variables, constants, function-letters and predicate letters (recall that the semantics of open formulas is postponed to Chaps. 7 and 8); (c) that the extension of $\mathfrak{I}$ in order to cover sentences requires a special definition of truth. Now we can go deeper into a deep contrast between **PC** and **FOL** as far as the semantic relations of sentences and their constituents are concerned. In the case of **PC**, it is enough to assume that atoms are true or false and proceed further by compositional recursive rules—starting with valuations of atomic sentences. In the case of **FOL**, the interpretation of **AL** does not determine truth-values of sentences (and the semantic interpretation of open formulas) in a recursive way, although it contributes essentially to this issue; details will be discussed in Chap. 8, Sect. 8.2). To anticipate perhaps the most crucial point: the full definition of truth, as offered by **STT**, appeals not only to metalogical concepts employed in the description of $\mathfrak{I}$ as it is formulated in the present chapter, but also to the notion of satisfaction.

Since quantifiers $\forall$ and $\exists$ (I neglect so-called numerical quantifiers, or instance 'there are exactly n objects') are undoubtedly logical constants, we can treat them just like propositional connectives—that is, as expressing functions of a kind (note, however, that quantifiers are not function symbols, that is, elements of the set **FUN**). The function corresponding to the universal quantifier maps it onto **U** (in other words, it shows the range of the universal quantifier), the function corresponding to the existential quantifier maps it onto its power set minus the empty set (it shows that the existential quantifier selects subsets and is related to the assumption that classical logic works only in non-empty universes). This treatment of quantifiers indicates their absolute sense (not depending on any particular interpretation) and motivates the fact that the reference to $\mathfrak{I}$ is omitted in (**Df11f**) and (**Df10g**). Accordingly, quantifiers have the same meaning under any interpretation. The status of identity leads to deep and serious problems. First of all, it is a (binary) predicate and its reference is established as any other predicate letter. In particular, (**Df11e**) correlates the identity sign with the binary relation of identity (equality, sameness, etc.). This suggests that identity belongs to extralogical concepts. Moreover, identity allows defining the mentioned numerical quantifiers. On the other hand, there is a very strong feeling of the absoluteness of identity, displayed by (**FOLA3**) (identity is the only predicate that applies to every object) and this circumstance suggests that identity belongs to the purely logical vocabulary. I took this path resulting in defining identity (see (**Df11d**)) without any reference to $\mathfrak{I}$. However, I admit that this solution is controversial. I return to this issue later in this section citing other arguments concerning the status of identity, as well as arguments for identifying of first-order logic with $\mathbf{FOL}_=$.

The concept of logical value (truth, falsehood; once again I do not enter into many-valued logics) employed in (**Df4**) has no application to (**Df11**). One might say that the issue here is only terminological, because there is no *a priori* reason to prevent a generalization of the concept of logical value by also considering individual objects, properties and relations as logical values; in fact, Frege (see Frege 1892) proposed such an extension. However, (**Df11**) has no counterpart of the function **v** from (**Df4**). Thus, the concept of logical value is somehow ambiguous. On the level of predicates, compositionality works as in **PC**. Consider the compound predicative form 'P_i and P_j', where both constituents are monadic (this example serves only an intuitive explanation, because we have no official notation for 'and' situated between predicate letters). Its value is given by

(8) $\mathbf{P_i} \cap \mathbf{P_j}$ (= $\mathbf{V}$ (Pj, $\mathfrak{I}$) $\cap$ $\mathbf{V}$ (Pj, $\mathfrak{I}$)).

Yet values of predicates, related to the interpretative sequence $\mathfrak{I}$, do not depend on values of terms. Assume that the value of P_i is $\{\mathbf{u_k}\}$ (=$\{\boldsymbol{P_i}\}$. Since $\mathbf{V}(P_i, \mathfrak{I})$ fixes the values in question, it has nothing to do with valuations of variables and constants. The predicate P_i can be supplemented by either a variable or by a constant. Thus, the general scheme of supplementation has the form $P_i(t_k)$. Consider now two valuations of terms

(9) (a) $\mathrm{V}_1: \quad t_1\left(\boldsymbol{u}_1^*\right), t_2\left(\boldsymbol{u}_2^*\right), t_3\left(\boldsymbol{u}_3^*\right), \ldots, t_k\left(\boldsymbol{u}_k^*\right), \ldots$

(b) $\mathrm{V}_2: \quad t_1\left(\boldsymbol{u}_1^*\right), t_2\left(\boldsymbol{u}_k^*\right), t_3\left(\boldsymbol{u}_3^*\right), \ldots, t_k\left(\boldsymbol{u}_2^*\right), \ldots$

Both valuations differ at places 2 and k, but this has no significance for the value of the predicate P_i. The set $\{\mathbf{u}_k\}$ always retains its value as long as the interpretative sequence remains unchanged. The set **U** itself as well all its subsets are fixed by $\mathfrak{I}$ independently of valuations of variables and constants—even independently of whether any constant is distinctive or not. This set plus individuals denoted by constants decide the world of predicate logic. Contrary to **PC**, for which linguistic resources allow defining only 20 different relations, **FOL** has no such limitation (this is displayed in a sense by the signature of pure **L**; see (**DG10**)). Thus the expressive power of **FOL** is much greater than that of **PC**, although not unlimited. We shall see in Sect. 5.4 that this circumstance constitutes the most critical feature in the discussion about the status of **FOL**.

(**DG11**) The above consideration is important for philosophy. Variables and constants look like the simples (the simple items of **L**). Accordingly, one can think of individual objects as ontological simples—because they function as values of logical simples. This view was adopted by Russell in his logic al atomism. However, no support for this view comes from logic. Since the function **V** for predicates does not depend on the valuation of terms, predicate letters and their values appear just as simple as terms and individuals. Moreover, every individual object belongs to some set. That means, ontologically speaking, that it has a property. Hence, there are no individuals without properties. The semantics of **FOL** does not tolerate bare particulars if they are thought of as devoid of properties. The situation described justifies to some extent the approach to set theory in which there are only sets. Although there can be other reasons—say empirical ones—for the view that individuals are prior to sets, but if formal ontology takes semantics for **FOL** as its conceptual pattern, the priority of individuals over sets raises serious doubts.►

(**DG12**) I adopt the view that $\mathfrak{I}$ is (I simplify the notation) the triple of the type <**AL**, $\mathfrak{R}$, **V**>. However, it is not the only possibility. Pogorzelski (see Pogorzelski 1994, p. 180) defines the interpretation as the pair <**U**, **V**>. Both perspectives make an explicit or implicit appeal to all three factors—namely **AL**, $\mathfrak{R}$ and **V**, and display to some extent certain problems pertaining the general concept of interpretation. We can think of interpretation as either an act or as a product. This ambiguity is preserved by the possible differing of **V**([…], $\mathfrak{I}$) and **V**([…], $\mathfrak{R}$). On the one side, they can express the process of interpreting a given expression, but, on the other side, one may treat them as information about how interpretation is fixed. I think that the understanding of interpretation as the triple <**AL**, $\mathfrak{R}$, $\mathfrak{I}$> is preferable, because this choice much better corresponds to the distinction of interpretation as act and as product, which is important from the philosophical point of view. The function **V** understood as a mapping displays the process of interpretation, but its values record interpretative results. Another useful device consists in introducing a distinction between the interpretandum (what is interpreted) and the interpretans (what interprets). It is very natural to assume that languages (or alphabets) are interpretanda. Thus, the content of $\mathfrak{I}$ shows what is interpreted, **AL**, what interprets, $\mathfrak{R}$, and how the interpretatndum is interpreted by the interpretans, **V**. If we agree that language (or alphabet) functions as the interpretandum, we become fairly strongly motivated to take the linguistic perspective (that is, beginning with languages) for defining interpretation as much more natural than the approach which considers $\mathfrak{R}$ as what is interpreted.►

We arrive at the point where the concept of model is to be introduced. Intuitively, a model of a language is a structure in which the sentences of a given **L** have logical values, that is, they can be conceived as true or false. Another way of speaking about models regards them as structures in which some subsets of **L** are true. Language as a whole cannot entirely consist of true sentences, because, in order to contain negation and satisfy (**BI**), it must have both true and sentences; I neglect the case in which **L** contains no negation, but has pairs of sentences that exclude each other (cannot be true together). Hence, if the model of a set **X** of sentences were to be defined as a structure in which all elements in **X** are true, languages in their entirety could not have models. Hence, models are structures corresponding to some subsets of **L**. Note, however, that if a structure is a model of a set of sentences, it also ascribes truth-values to other sentences, because negations of the members of **X** are false in that structure. The exact definition of model is as follows:

(**Df12**) Let **X** be set of sentences. A model of a set **X** is a structure $\mathbf{M} = <\mathbf{U}, \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \ldots, \mathbf{P}_{1(1)}, \mathbf{P}_{1(2)}, \ldots, \mathbf{P}_{2(1)}, \mathbf{P}_{2(2)}, \ldots$, such that for every $A \in \mathbf{X}$,
$\mathbf{M} \models A$ (means: 'A is true in **M**'). Briefly: $\mathbf{M}(\mathbf{X}) \Leftrightarrow \forall A \in \mathbf{X}, \mathbf{M} \models A$. If it is not the case that $\mathbf{M} \models A$, we write $\mathbf{M} \dashv A$.

It is easily to see that **M** and $\mathfrak{R}$ (the interpretative structure) are not distinguishable from the point of view of their content, because the right sides of the corresponding equalities are exactly the same. The actual difference between models and interpretations shows up when truth-conditions become involved. This fact constitutes perhaps the main reason for adopting the first perspective in explaining the concept of interpretation and, in consequence, for distinguishing interpretative structures from models. Moreover, interpretations concern **AL** directly and **L**, considered as a set of sentences indirectly. Yet models are objects in which we consider just sentences directly, but their constituents indirectly. This explains why truth appears as directly associated with model, but mediated by an interpretation. It is one of the most important features of **STT**. The term 'semi-model' can be applied to interpretative structures as well, but is only a linguistic convention and has no substantial importance.

As I already noted, I shall not define the concept of truth in its full generality in this chapter. However, in order to point out some important intuitions, it seems proper to give a definition for a special case—related to so-called canonical models. They are structures in which every object has its own individual name. The standard model of the arithmetic of natural numbers is an example here, because every numeral functions as the proper name of a corresponding number. On the other hand, the model of the arithmetic of real numbers cannot be canonical owing to the uncountability of reals. If a model **M** is canonical, we do not need to consider open formulas, because for any formula of the type $P_{i(j)}(x_{1(j)},\ldots, x_{i(j)})$, we have a corresponding sentence $P_{i(j)}(t_{1(j)}, \ldots, t_{i(j)})$ in which no term is a variable. For simplicity, I limit the task to unary predicates (and identity) and neglect function letters; the generalization to the fully equipped first-order **L** is straightforward. The definition is as follows (I call it 'the special definition of truth'):

(**SDfVER**) (**a**) $\mathbf{M} \models^{\mathfrak{I}} P_{1(j)}(t) \Leftrightarrow \mathbf{t} \in \mathbf{P_{1(j)}}$;
(**b**) $\mathbf{M} \models^{\mathfrak{I}} (t_i = t_j) \Leftrightarrow \mathbf{Id}(\mathbf{t_i}, \mathbf{t_j})$;
(**c**) $\mathbf{M} \models^{\mathfrak{I}} \neg A \Leftrightarrow \mathbf{M}$ not- $\models^{\mathfrak{I}} A$;
(**d**) $\mathbf{M} \models^{\mathfrak{I}} A \wedge B \Leftrightarrow \mathbf{M} \models^{\mathfrak{I}} A$ and $\mathbf{M} \models^{\mathfrak{I}} B$;
(**e**) $\mathbf{M} \models^{\mathfrak{I}} A \vee B \Leftrightarrow \mathbf{M} \models^{\mathfrak{I}} A$ or $\mathbf{\Omega} \models^{\mathfrak{I}} B$;
(**f**) $\mathbf{M} \models^{\mathfrak{I}} A \Rightarrow B \Leftrightarrow \mathbf{M}$ not- $\models^{\mathfrak{I}} A$ or $\mathbf{M} \models^{\mathfrak{I}} B$;
(**g**) $\mathbf{M} \models^{\mathfrak{I}} A \Leftrightarrow B$ iff if $\mathbf{M} \models^{\mathfrak{I}} A$ and $\mathbf{M} \models^{\mathfrak{I}} B$ or $\mathbf{M}$ not- $\models^{\mathfrak{I}} A$ and $\mathbf{M}$ not- $\models^{\mathfrak{I}} A$
(**h**) $\mathbf{M} \models^{\mathfrak{I}} \forall vA(v) \Leftrightarrow \mathbf{M} \models^{\mathfrak{I}} A(v)$, for every formula $A(v/t)$, if t is not a variable;
(**i**) $\mathbf{M} \models^{\mathfrak{I}} \exists vA(v) \Leftrightarrow \mathbf{M} \models^{\mathfrak{I}} A(v)$, for some formula $A(v/t)$, if t is not a variable.

The particular points say that: (**a**) a formula of the type $P_{1(j)}(t)$ is true, provided that the object named by t belong to the set $\mathbf{P_{1(j)}}$; (**b**) an identity sentence is true, provided that denotations of its nominal constituents are identical; (**c**) negation of a sentence A is true, provided that A is not true (note that this condition does not express bivalence without additional constraints); (**d**), (**e**), (**f**) and (**g**) repeat the semantic rules of **PC** conjunction, disjunction, implication and equivalence as extended to **FOL**; (**h**) the universal sentence $\forall vA(v)$ is true, provided that the formula $A(v)$ becomes true under all substitutions of the variable v by an individual constant; (**i**) the existential sentence $\exists vA(v)$ is true, provided that the formula $A(v)$ becomes true under some (at least one) substitution of the variable v by an individual constant. All points are implicitly relativized to $\mathfrak{I}$; the importance of this restriction will be explained later (see Chap. 8, Sect. 8.2). Semantics based on the concept of model is called model-theoretical, due to the treatment of truth conditions for sentences.

The label 'model-theoretic semantics' clarifies many issues, especially topics related to the concept of identity. A general form of identity sentences is given by the following formula

(9) $t_i = t_j$,

where t_i and t_j are terms. Assume that $\mathbf{v}(t_i, \mathfrak{I}) = \boldsymbol{u_i}$ and $\mathbf{v}(t_j, \mathfrak{I}) = \boldsymbol{u_j}$. Thus, we have

(10) $\mathbf{M} \models^{\mathfrak{I}} (t_i = t_j) \Leftrightarrow \mathbf{Id}(\mathbf{v}(t_i, \mathfrak{I}), \mathbf{v}(t_j, \mathfrak{I})) \Leftrightarrow \boldsymbol{u_i} = \boldsymbol{u_j}$.

Using the example from (**DG8**), we say that the sentence:

(11) Plato = Aristocles is true, because in the historical world (taken as $\mathbf{\Omega}$)
(12) $\mathbf{Id}(\mathbf{v}$('Plato'), $\mathbf{v}$('Aristocles')).

Identity is troublesome not only from the point of view of its logical status, expressed in the question "Is it a logical constant or not", but also because of its puzzling paradoxes. Frege (see Frege 1891) observed the following simple problem. Consider (the examples are adapted for (11))

(13) Plato = Plato.

Now, if we accept (13), substituting one occurrence of 'Plato' by 'Aristocles' in (13), gives (11). However, although (12) is a triviality as every formula of the type $a = a$,, (11) provides a substantial historical information. In fact, nobody worries about 'Plato = Plato", but (11) must be supported by historical sources. A given person can know (13) (he or she eventually asks who was or is Plato), but yet not realize that (11) holds; for example, my guess is that even most philosophers has no idea that 'Aristocles' was Plato's original name. The problem is that a logical operation that seems to be fully admissible (the substitution of terms referring to identicals), leads to evident troubles. Frege solved the problem by his famous distinction of sense and reference: (11) and (13) refer to the same (for Frege, to the True), both radically differ in their senses.

There is no doubt that (11) and (13) differ in regard to their corresponding senses (or meanings, if some prefers this way of speaking). But I am still convinced that Frege's puzzle is artificial. To begin with, let me observe that putting 'Aristocles' in the place of 'Plato' (or performing the reverse operation) could not be regarded as substitution in the strict logical sense, because we only substitute for free variables, not for constants. To be logically correct, we should rather say that the term 'Plato' is replaced by the term 'Aristocles'. However, there is no rule of replacement among the rules of inference of **FOL**. The axiom (**FOLA4**) cannot be used because it assumes identity. We have at our disposal only (13), but we need to prove (11). Assume that the full version of the Leibniz law, namely the formula:

(**LL**) $(v_1 = w_1) \wedge \ldots \wedge (v_n = w_n) \Leftrightarrow (P_{n(i)}(v_{1(i)}, \ldots, v_{n(i)}) \Leftrightarrow P_{n(i)}(w_{1(i)}, \ldots, w_{n(i)}))$,

is added to the axioms of **FOL**. Now, we can apparently argue in the following way. Firstly, we observe

(14) Plato is Plato $\Leftrightarrow$ Plato is Aristocles.

Then, we conclude, employing (**LL**), that Plato is just identical with Aristocles. Leaving aside the problematic character of (**LL**) (more precisely, the implication from right to left, because from the left to right is obvious), the whole inference is suspicious. In fact, (14) has no derivation proceeding by purely logical means, unless (11) is assumed as a logical truth. In particular, this assumption would result in the circularity. The actual situation seems be this. There are no other purely logical identity sentences than provable from (**FOL3**) and (**FOL4**) by (**MP**), (**Ad**$\forall$) and (**Ad**$\exists$). In order to obtain (11), we need to use (11), which, in general, is an extralogical premise, except for the case $\mathbf{V}(t) = \mathbf{V}(t)$. Even in mathematics, the identity of 4 and $2 \cdot 2$ is not purely logical. Since we assume that terms are valued by the valuation function, it is nothing strange that (11) provides a substantial historical information, but (13) is just a tautology.

Definition (**Df5**) is easily adapted for **FOL**. We have

(**Df13**) (**a**) A is FOL-satisfiable if and only if for some $\mathbf{M}$, $A \in \mathbf{X}$, $\mathbf{M} \models \boldsymbol{A}$;
(**b**) $\mathbf{X}$ is **FOL**-satisfiable if and only if for every $A \in \mathbf{X}$, A is **FOL**–satisfiable;
(**c**) A is a **FOL**-tautology ($A \in \mathrm{TAU}_{\mathbf{FOL}}$, $\models_{\mathrm{FOL}} \boldsymbol{A}$) if and only if for every $\mathbf{M} \models A$;
(**d**) A is a **FOL**-semantic consequence of $\mathbf{X}$ ($\mathbf{X} \models_{\mathbf{FOL}} A$) if and only if $\mathbf{M} \models \mathbf{X}$, provided $\mathbf{M} \models A$;
(**e**) A is a **FOL**-contradiction ($A \in \mathbf{CTL}_{\mathbf{FOL}}$, $_{\mathbf{FO:}}\!\!=\!\!| \boldsymbol{A}$) if and only if A is not **FOL**-satisfiable, that if for any $\mathbf{M}$, it is not the case that $\mathrm{M} \models A$.

Since **FOL**, as any logical system, codifies truth-preserving inferences, the restrictions imposed on inferential rules prescribed for **FOL** find their demonstrative justification by pointing out their role in securing that truth is always transmitted from premises to conclusions. Why can the variable v not be free in the antecedent of the upper formula in (**Ad**$\forall$), that is, the formula A? This can be answered by two examples related to (**Df8**) and (**Df9**). The formula

(15) $P(x) \Rightarrow P(x)$,

is a tautology, because it is an instance of the formula $A \Rightarrow A$, which is universally valid in **PC**. Applying (**Ad**$\forall$) leads to

(16) $P(x) \Rightarrow \forall x P(x)$.

Since the occurrence of the variable x is free in the antecedent, we can substitute a term for x, say the constant a. Thus, we receive

(17) $P(a) \Rightarrow \forall x P(x)$,

which is not logically valid, because the sentence 'something is P' does not imply 'everything is P', unless we restrict our considerations to an universe that contains exactly one element. Hence, it can happen that (15) is true, but (17) becomes false. Observe that the application of the rule (**Ad**$\forall$) to (15) plays the critical role in the above inference. Consider now the formulas

(18) (a) $x + 1 > 2$;

(b) $x > 1$,

assuming that the variable x ranges over natural numbers. The implication

(19) $(x + 1 > 2) \Rightarrow (x > 1)$,

is true for all natural numbers greater than 1 (recall that (**Df9**) does not require that all lines of a proof must be tautologies). Applying (**Ad**$\forall$), we obtain

(20) $(x + 1 > 2) \Rightarrow \forall x(x > 1)$.

Substituting the numeral 3 for x gives

(21) $(3 > 2) \Rightarrow \forall x(x > 1)$,

which is false. However, if we start with

(22) $(3 > 2) \Rightarrow (x > 1)$,

this problem disappears, because it is a false sentence (it is enough to substitute 1 for x). Thus, (22) does not destroy the principle that logic is truth-preserving. This example shows how semantics illuminates the issues of syntax.

5.3 Selected Concepts and Theorems of Metalogic (Metamathematics)

Let **X** be an arbitrary set of sentences. We define

(**Df14**) $A \in Cn\mathbf{X}$ if and only if $\mathbf{X} \vdash A$

Thus, A is a consequence of **X** if and only if A has a proof on the basis of (is derivable from) **X**. Although Cn (the consequence operation) and $\vdash$ (the consequence relation) are mutually interdefinable, the two are categorically different. If **L** is a language understood as a set of formulas. Cn maps $\mathbf{2^L}$ onto $\mathbf{2^L}$ (sets of formulas onto sets of formulas), but the consequence relation acts on the subsets of $\mathbf{2^L} \times \mathbf{L}$, that is, links sets of formulas with single formulas. By definition, properties of the operation Cn are determined by (**Df13**). Assume, however, that we have a very vague idea about inferring sentences from other sentences. How many mappings of the type Cn (or $\vdash$) are there? The answer says that we have uncountably many of different consequence operations. Hence, it is very important to discriminate "good" consequence operations (relations). (**Df13**) points out one way of doing that. Another path is due to Tarski (see Tarski 1930, Borkowski 1991) and consists in axiomatizing properties of Cn. The set of axioms related to **FOL** (= $\mathbf{FOL_=}$) is as follows:

(**CnA1**) $\varnothing \leq \mathbf{L} \leq \aleph_{0;}$
(**CnA2**) $\mathbf{X} \subseteq Cn\mathbf{X}$;
(**CnA3**) $\mathbf{X} \subseteq \mathbf{Y} \Rightarrow Cn\mathbf{X} \subseteq Cn\mathbf{Y}$;
(**CnA4**) $CnCn\mathbf{X} = Cn\mathbf{X}$;
(**CnA5**) $A \in Cn\mathbf{X} \Rightarrow \exists \mathrm{Y} \subseteq \mathbf{X} \wedge \mathbf{Y} \in \mathbf{FIN} \wedge (A \in Cn\mathbf{Y})$;
(**CnA6**) $(A \Rightarrow B) \in Cn\mathbf{X} \Rightarrow B \in Cn(\mathbf{X} \cup \{A\})$;
(**CnA7**) $B \in Cn(\mathbf{X} \cup \{A\}) \Rightarrow (A \Rightarrow B) \in Cn\mathbf{X}$;
(**CnA8**) $Cn\{A, \neg A\} = \mathbf{L}$;
(**CnA9**) $Cn\{A\} \cap Cn\{\neg A\} = Cn\varnothing$;
(**CnA10**) $A(v/t) \in Cn\{\forall vA(v)\}$, if the term t is substitutable for v;

(**CnA11**) $A \in CnX \Rightarrow \forall vA(v) \in CnX$, if v is not free in B, for any $B \in X$;
(**CnA12**) $\forall t_i(t_i = t_i) \in Cn\varnothing$;
(**CnA13**) $((s_1 = t_1) \wedge \ldots \wedge (s_n = t_n)\ldots) \Rightarrow A(t_1,\ldots, t_n) \in Cn\{A(s_1,\ldots, s_n\})$.

The set {(**CnA1**)–(**CnA13**)} can be divided into two groups. The first group includes (**CnA1**)–(**CnA5**) as general axioms for *Cn*. (**CnA1**) says that the cardinality of **L** is at most denumerably infinite, (**CnA2**)—that any set is a subset of the set of its consequences, (**CnA3**) established the monotonicity of *Cn*, (**CnA4**) its idempotency, (**CnA5**) states the finiteness condition, which means that if something belongs to *Cn***X**, it belongs to the set of consequences of a finite subset of **X** (the notation **X** $\in$ **FIN** means '**X** is a finite set'). In other words: every inference is finitary, that is, performable on the basis of a finite set of premises and, according to the character of the rules, has finite length. (**CnA1**)–(**CnA5**) do not provide any logic in its usual sense, because they do not generate any rules of inference. They rather characterize *Cn* as a kind of closure operator. The logical machinery associated with *Cn* is encapsulated by the rest of the axioms (related to logic based on negation, implication and the universal quantifier, that is, **PC** and **FOL**). (**CnA6**) formulates the deduction theorem (see (**DT**) below), but if it is to be applied to predicate logic, we must assume that A, B $\in$ **SEN**; (**CnA7**) is (**MP**), (**CnA8**)–(**CnA9**) characterize negation, (**CnA10**)–(**CnA11**) are related to the universal quantifier, and (**CnA12**)–(**CnA13**) deal with identity.

(**DG13**) What does it mean that *Cn* as axiomatized by (**CnA1**)–(**CnA13**) is good? Two intuitions are captured by the positive answer. Firstly, *Cn* preserves truth, and, secondly, it is associated with classical logic. In Chap. 4, Sect. 4.8 and (**DG15IV**), an example of reasoning that preserves falsehood was given and commented on. Dual logic (I omit other logics preserving in which **0** is can be the distinctive value)

(**Df*dCn***) $A \in dCn$**X** if and only if $\exists$**Y**(**Y** $\subseteq$ **X**) and $\cap$ ($Cn\{B\}$: $B \in$ **Y**) $\subseteq Cn\{A\}$
Roughly speaking, A is a dual consequence of **X** if its dual counterpart is a *Cn*-consequence of a dualized subset of **X** (I omit the definition of duality). *dCn* can be axiomatized and has properties analogous to those of *Cn.* The motivation for choosing the latter consequence is simply its relation to truth (see (**DG15IV**) for a justification of this view). However, the dual logic is perfectly correct from a formal point of view.

Why is classical logic preferred in this book? I might answer that it is for its close connection with **STT**. However, one might object that this pragmatic answer does not suffice from a general philosophical point of view, because we have several mutually rival logical systems—for instance, fuzzy, many-valued or intuitionistic. Are they really rival, or complementary? Look at axioms (**PCA1**)–(**PCA15**). These characterize the classical system, but if we drop (**PCA15**) we obtain intuitionistic (propositional) logic. This fact motivates the view that classical logic includes intuitionistic logic as a part and, thereby, the former is more general than the latter. However, defenders of intuitionism as the correct philosophy of logic and mathematics say that classical logic should be rejected and the intuitionistic adopted as logical foundation. The most radical intuitionists even say that they do not understand classical logic. This means that both systems are incomparable. Yet other problems arise, when many-valued logic is taken into account because it intersects with the classical system. In fact, we have two general positions, namely logical pluralism (there are many parallel logics, which have different applications and none of them can be considered as *the* universal logic) and logical absolutism (there is one proper logical system, and other logics are either its species or should not be called logics at all). Since this book is not an essay in the philosophy of logic, I will not go beyond these very elementary remarks (see Haack 1978, Chap. 9, Beall, Restall 2006, Shapiro 2014 for more extensive treatments). However, some particular problems will be discussed in conjunction with **STT**.►

Having *Cn* (or $\vdash$, but I choose *Cn*, except some cases), we can simply define several important concepts used in metalogical and metamathematical investigations (**AR** refers to the arithmetic of natural numbers). It is done by:

(**Df15**) (**a**) **X** is a deductive system ($\mathbf{X} \in \mathbf{SYS}$) if and only if $Cn\mathbf{X} \subseteq \mathbf{X}$;
(**b**) **X** is (finitely, recursively) axiomatizable ($\mathbf{X} \in \mathbf{AX}$) if and only for some (finite, recursive) set $\mathbf{Y} \subseteq \mathbf{X}$, $Cn\ \mathbf{Y} = \mathbf{X}$;
(**c**) **Y** is an independent set of axioms for **X** if and only (c_1) $\mathbf{Y} \subseteq \mathbf{X}$; ($c_2$) $Cn\mathbf{Y} = \mathbf{X}$; ($c_3$) for any $A \in \mathbf{Y}$, $A \notin Cn(\mathbf{Y} - \{A\}$;
(**d**) **X** is absolutely consistent ($\mathbf{X} \in \mathbf{CONS_A}$) if and only if $Cn\mathbf{X}$ ^{1} **L**;
(**e**) **X** is negation consistent ($\mathbf{X} \in \mathbf{CONS_N}$) if and only if for any A, $A \wedge \neg A \notin Cn\mathbf{X}$;
(**f**) **X** is syntactically complete ($\mathbf{X} \in \mathbf{COM_{SYN}}$) if and only if for any A, A or $\neg A \in Cn\mathbf{X}$;
(**g**) **X** is strongly semantically complete ($\mathbf{X} \in \mathbf{COM_{SSEM}}$) if and only if for any A, $\mathbf{X} \vdash A$ iff $\mathbf{X} \models A$;
(**h**) **X** is weakly semantic complete ($\mathbf{COM_{WSEM}}$) if and only if $\vdash A \Leftrightarrow \models A$;
(**i**) **X** is Post-complete or maximally consistent ($\mathbf{X} \in \mathbf{COM_P}$) if and only if $\mathbf{X} \in \mathbf{CON_A}$ and for any $A \notin \mathbf{X}$, $Cn\{\mathbf{X} \cup \{A\} = \mathbf{L}$;
(**j**) X is compact for consistency (**COMP**) **X** is consistent if and only if every finite subset of it is consistent;
(**k**) **X** is decidable ($\mathbf{X} \in \mathbf{DEC}$) if and only if there is an algorithmic (mechanical) procedure that solves the problem of whether for any $A \in \mathbf{X}$, $A \in \mathrm{X}^{\mathbf{Cn}}$ (the notation $A \in \mathbf{X}^{\mathbf{Cn}}$ means 'A is a theorem of **X**; it is equivalent to $A \in Cn\mathbf{X}$);
(**l**) **X** (provided that **AR** can be formalized in it) is ω-consistent ($\mathbf{X} \in \boldsymbol{\omega}\mathbf{CONS}$) with respect to the sequence $t_1, \ldots, t_k, \ldots$ of its terms if and only for any $A(x) \in \mathbf{X}$ (x is free in $A(x)$), $\forall t \in \{ t_1, t_2, t_3, \ldots, t_k, \ldots\} A(x/ t_k) \in \mathbf{X} \Rightarrow \neg\exists x \neg A(x) \in \mathbf{X}$).

According to (**Df15a**), **X** is a deductive system if the consequences of **X** are contained in it. Due to (**CnA2**) and the definition of a deductive system, we can also say that $\mathbf{X} \in \mathbf{SYS}$ if and only if $\mathbf{X} = Cn\mathbf{X}$, that is, a set is a deductive system if it is equal to the set of its consequences. (**Df15b**) deals with the concept of axiomatizability. A set is axiomatizable if it has a subset that produces all its consequences; the second set is an axiomatic basis of (or an axiom system for) the first set. Since every set is a subset of itself, every set forms own axiomatization. This, however, is too trivial of a case to be interesting. Thus, **Y** is a non-trivial axiomatic basis for **X** if and only if $\mathbf{Y} \neq \mathbf{X}$. If **Y** is finite, then **X** is finitely axiomatizable. **PC** in the version given above is not finitely axiomatizable because every axiom-scheme (**PCA1**)–(**PCA15**) represents infinitely many formulas. However, if these axioms are replaced by concrete formulas—for example, the axiom (**PCA1**) by the formula $p_1 \Rightarrow (p_2 \Rightarrow p_3)$—the axiomatic basis of **PC** becomes finite; the same applies to **FOL**. This situation is a good example of how metalogical properties depend on syntax (see also the remarks about Post-completeness, below). Yet one can say that if an infinitely axiomatizable system can be replaced by a finitely axiomatizable, one the difference between them is not very essential. In fact, there is common agreement that **FOL** is the same logical system in both versions. (**Df15c**) (change the succession of comments slightly) concerns independence as a property of axiom-systems. An axiom-system is independent if its elements—that is, particular axioms—are not deducible from others. Let **Ax** be an axiom-system and let $A \in \mathbf{Ax}$. In order to prove that A is independent of the rest of **Ax**, it is sufficient to find two models **M** and **M′** such that $\mathbf{M} \models A$ and $\mathbf{M}' \models \mathbf{Ax}'$, where **Ax′** arises from **Ax** by eliminating A, and $\mathbf{M}' \Dashv A$. Here we have an application of the concept of model. The given axiomatizations of **PC** and **FOL** are independent.

A system can be recursively axiomatized or not. The concept of recursivity is too complicated to be fully explained in the present book (see Murawski 1999). Roughly speaking, recursive functions are functions form the set of natural numbers to the same set. This means that these functions have natural numbers as arguments and values; for example, addition and multiplication are just recursive. The general idea is that recursive functions are effectively (via algorithms) calculable. The identification of an intuitive concept of effective calculability and the exact concept of recursivity is called the Church thesis:

(**ChT**) The class of effective calculable functions = the class of recursive functions.

Since every recursive function is effectively calculable, the real content of (**ChT**) can be expressed as the claim that every calculable function is recursive. Now, a set is recursive if and only if its elements are effectively calculable. Both axiomatic bases of **FOL** are recursive, although one of them is not finite. On the other hand, the set of theorems of **FOL** and the set of theorems of arithmetic of natural numbers are not recursive because both theories are undecidable (see Chap. 8, Sect. 8.4). All finite sets, **AL**, **L**, the sets of proofs in the sense of (**Df2**) and (**Df3**) (or (**Df8**) and (**Df9**)) are further examples of recursive sets.

A set **X** is absolutely consistent (condition (**Df15d**)) if and only if something does not belong to its logical consequences. The negation consistency of **X** (condition (**Df15e**) amounts to the property that no pair of the type $\{A, \neg A\}$ is among **X**-consequences. If negation belongs to **L**, both concepts of consistency are equivalent. Consistency is perhaps the most important metalogical concept, because it refers to a mandatory property of deductive systems. One distinguishes relative and absolute proofs of consistency. The former consists in proving the following statement about a system **X**:

(23) If **X'** is consistent, then **X** is consistent,

provided that **X** is interpretable. If (23) is proved for a given set **X**, we say that consistency of **X** was proved relatively (with respect) to the consistency of **X′**. In the 19th century, Eugenio Beltrami demonstrated that non-Euclidean geometry is consistent relative to Euclidean geometry via interpreting the latter in the former (I omit a general notion of interpretability). If **X**–consistency is provable without assuming that something else (except logic) is consistent, the absolute consistency of **X** is available. Hence, logic—as assumed in every proof—must be consistent before anything else could be consistent.

The conditions (**Df15f**)–(**Df15i**) review three concepts of completeness. Syntactic completeness of a given system **X** means that either A belongs to **X**-theorems or $\neg A$ belong to them. Weak semantic completeness is a special case of strong completeness, because it is enough to put empty set in the place of **X** in order to obtain (**Df15g**) from (**Df15h**). However, the concept of weak completeness is not trivial (see Sect. 5.4 on the definition of logic). In general, the completeness property, if possessed by a system **X**, indicates that its syntactic equipment is equivalent with its semantic machinery. Completeness can be split into two complementary properties—namely, completeness proper (every valid sentence is provable), and soudness or correctness (every provable sentence is valid); in the case of logical rules, soudness consists in truth-preservation. If a system **X** is Post-complete (condition (**Df15i**), it is maximal in the sense than no sentence can be added to it without producing inconsistency. Compactness (see (**Df15j**) indicates that consistency “moves” via finite sets from bottom to top as well as in the reverse direction; the first direction is more important. Decidability (see (**Df15k**) concerns the problem of whether the property ‘being a theorem’ of’ is checkable in an algorithmic, that is, recursive way. A sensational discovery was the proof that decidability and provability are different properties. It can be demonstrated that A is provable on the basis of **X** without giving an algorithm for checking that. Condition (**Df15l**) describes ω-consistency, the property that is applicable to systems in which the arithmetic of natural numbers is expressible. A system has this property provided that if every monadic formula $A(t_k)$ obtained by substituting the term t_k for the variable x the term t_k, is provable in this system, the formula $\exists x \neg A(x)$ is not provable. The concept of ω-consistency is stronger than consistency. This means that there are systems that are consistent but are just ω-inconsistent (see Chap. 8, Sect. 8.3). In general, $\mathbf{CONS_A}$, $\mathbf{COM_{SYN}}$, **DEC** and **ωCONS** are syntactic concepts, whereas **COMP** and $\mathbf{COM_P}$ are considered both from a semantic (model-theoretic) and syntactic point of view; the status of $\mathbf{COM_{SSEM}}$ and $\mathbf{COM_{WSEM}}$ was already explained.

The next two statements summarize the properties of **PC** and **FOL**:

(24) **PC** is a deductive system, which

(a) is finitely or recursively axiomatizable;

(b) has several independent axiomatic bases;

(c) is not syntactically complete;

(d) is strongly semantically complete ($\mathbf{X} \vdash_{\mathbf{PC}} A \Leftrightarrow \mathbf{X} \models_{\mathbf{PC}} A$);

(e) is weakly semantically complete ($\varnothing \vdash_{\mathbf{PC}} A \Leftrightarrow \varnothing \models_{\mathbf{PC}} A$);

(f) is compact;

(g) is Post-complete (strictly speaking, it concerns **PC** as formalized by concrete formulas, but a similar property can be defined for its codification by schemata);

(h) is decidable.

(25) **FOL** is a deductive system, which

(a) is finitely or recursively axiomatizable;

(b) has several independent axiomatic bases;

(c) is not syntactically complete;

(d) is strongly semantically complete ($\mathbf{X} \vdash_{\mathbf{PC}} A \Leftrightarrow \mathbf{X} \models_{\mathbf{PC}}$);

(e) is weakly semantically complete ($\varnothing \vdash_{\mathbf{PC}} A \Leftrightarrow \varnothing \models_{\mathbf{PC}}$);

(f) is compact;

(g) is not Post–complete;

(h) is undecidable.

(**DG14**) Any set of true sentences is a deductive system, because the consequences of true sentences are also true. However, false sentences do not form a deductive system, because consequence of falsehoods can be true or false. The situation changes when we operate with *dCn*. We can define a deductive system related to the dual consequence by (a) $\mathbf{X} \in \mathbf{dSYS}$ if and only if $\mathbf{X} = dCn\mathbf{X}$. All metalogical concepts and theorems discussed above and below can be easily adapted to dual logic.►

The inspection of (24) and (25) shows that the two parts of elementary logic have different properties. They share features (a)–(g), but differ with respect to (f) and (g). That **PC** and **FOL** are not syntactically complete can be easily demonstrated by showing that no atomic sentence (represented by a sentential variable) and its negation are theorems of **PC**, and no formula of the type $P_{1(i)}(x)$ and $\neg P_{1(i)}(x)$ belong to theorems of **FOL**. Consider **PC** as the set of tautologies. One can show that $\mathbf{PC} \cup \{A\}$, where A is not tautology, is inconsistent. It follows from Post-completeness Consider now **FOL** as the set of tautologies. The system **FOL** ∪ {'there is exactly one individual'} is consistent. The truth-table method provides an algorithm for decided whether an arbitrary formula A is (or not) a theorem of **PC**. No such procedure is available for **FOL**. Thus, although we know that every tautology is provable, we have no method to show that the class of tautologies is recursive.

The completeness theorem also has another form the Gödel–Malcev theorem), namely

(26) A set **X** is consistent if and only if has a model.

The completeness property in the sense of (26) is equivalent to the strong completeness. Having (26) we can restate compactness by

(27) A set **X** has a model iff its every finite subset has a model.

I will now list four important theorems that will be used in further considerations:

(28) Every consistent set of sentences has a maximal consistent extension;
(29) A set of first-order sentences has a denumerable model if and only if it has a model of arbitrary cardinality;
(30) If a system with classical connectives and a denumerable alphabet is compact or complete and satisfies (28) (more strictly: the then-part of (28)), it is equivalent to **FOL**;
(31) **FOL** does not distinguish any extralogical concept—if something can be proved in **FOL** about a property or relation not belonging to pure logic, it can also be proved about any other extralogical property or relation.

Theorem (28) due to Adolf Lindenbaum, says that for every consistent set of sentences, exists a maximally consistent oversystem, that is, a set of sentences which does not tolerate any extension without producing inconsistency. The statement (29) captures the Löwenheim–Skolem–Tarski theorem. More precisely, its then-part (if a first-order set of sentences, that is, formulated in the language of **FOL**, has a model of arbitrary infinite cardinality, then it has a denumerable model; it is called the downward Löwenheim–Skolem theorem) was proved by Leopold Löwenheim and improved by Thoralf Skolem, and the if-part (if a first-order set of sentences has a denumerable model, it has a model of arbitrary cardinality; it is sometimes called the upward Löwenheim–Skolem theorem) is due to Tarski. (30) reproduces the content of the Lindsröm theorem, which characterizes **FOL** and says that it is the only logic which is complete or compact and satisfies the Löwenheim–Skolem theorem (or has the Löwenheim–Skolem property) (see Flum 1985 for an extensive treatment). Finally, (31) displays the fact that logic is entirely neutral with respect to extralogical matters. This theorem holds for pure as well as applied **FOL**, provided that no extralogical axioms are added.

Metalogic helps us to solve (or at least to illuminate) the controversial issue of the status of identity as somehow placed between propositional connectives and quantifiers being purely logical items and extralogical concepts. The main reason in order to include the identity predicate in the list of logical constants is that $\mathbf{FOL}_=$ satisfies the main metalogical theorems concerning elementary logic, namely (26)–

(31). This suggests that identity behaves like logical constants of **PC** and **FOL** (without identity) On the other hand, since identity makes it possible to define numerical quantifiers (see above), like (to remind) ‘there are exactly two’, ‘there are exactly three’, etc. (for an arbitrary natural number *n*), it seems to introduce extralogical contents to logic, and, should thereby not be considered as a logical notion. This controversy can be settled only by a decision, and my choice is to include identity among logical constants. Thus, **FOL** refers to first-order (classical) logic with identity. Nevertheless, the symbol **FOL** is sometimes used in further considerations as referring to first-order logic without identity, but I hope that associated comments will prevent misunderstandings. However, it should be added that status of identity suggests that the borderline between the logical and the extralogical is vague to some extent.

(**DG15**) $\mathbf{FOL}_{=}$ does not satisfy all metatheorems valid for first-order logic without identity. In particular, adding identity to the vocabulary of first-order logic results that the inflation and deflation theorems do not hold for **FOL**. The inflation theorem says that if a formula is valid in a universe with *n* elements, it is also valid in every greater universe; the deflation theorem says that if a formula is valid in a universe with *n* elements, it is also valid in every smaller universe. Since having identity we can define numerical quantifiers of the type ‘there are exactly *n* objects’, the formula ‘there are exactly *n* objects having property *P*’ provides a counterexample for inflation and deflation theorems. Three remarks are in order here. Firstly, it happens that subsystems satisfy given theorems which do not hold for their oversystems. For example, monadic predicate calculus, that is, a subsystem of **FOL** that admits only monadic predicates is decidable, contrary to the full first-order logic; **PC** is Post-complete, but **FOL** is not. Secondly, the Löwenheim–Skolem–Tarski theorem provides a generalization of inflation and deflation theorems that pertain to models with infinite universes. Thirdly, inflation and deflation theorems do not belong to the main results in metalogic. This suggests that they are not very relevant for defining logic.►

5.4 How to Define Logic?

This section focuses on the concept of logic. I consider the universality property as the most essential attribute of logic. Although none of the mentioned metalogical results says anything directly about the universality property and its aspects as defined in Sect. 5.1, metalogic can be used for illuminating some points. In particular, formal analysis provides grounds to consider (**Un1**)–(**Un3**) as equivalent, and contrasted with (**Un4**).

I shall begin with the way to define logic conceived as a deductive manual. Intuitively speaking, such manuals provide instructions on how to prove some propositions on the basis of others adopted as premises. It is done by means of inference rules; for example, (**MP**) informs us that it is logically acceptable to pass from A and $A \Rightarrow B$ as premises to B as the conclusion. The inference rules are hidden in *Cn.* This suggests using, as starting point, the concept of consequence operation as axiomatized in Section 3 (see Wójcicki 1988 for an extensive and detailed analysis of logical calculi via *Cn*). How to define logic *via Cn*? Having the deduction theorem, we say that **LOG** is identified as $Cn\varnothing$. More formally we have:

(**DfLOG1**) $A \in$ **LOG** $\Leftrightarrow A \in Cn\varnothing$, or, equivalently **LOG** = $Cn\varnothing$.

At first sight this definition looks artificial at first sight; clearly, here the empty set looks here like a convenient metaphor. In particular, one might argue that we can derive something from the empty set only because of the logical machinery is already incorporated into *Cn*. Otherwise speaking, we tacitly assumed that axioms for *Cn* have a certain logical content. Hence, the question arises how to justify that stipulations (**CnA1**)–(**CnA13**) about the consequence operation are proper for logic. As far as the general axioms are concerned, we can for instance drop the requirement of monotonicity (it leads to non-monotonic logics used in computer science) or finiteness in order to obtain infinitary logics, that is, logics with infinitely long formulas. Hence, any definition of logic *via* the consequence operation needs additional justification.

(**CnA1**) and (**CnA5**) are closely related to the human faculties in doing inferences. A possible defence of these axioms consists in pointing out that our inferential performances have a finitary character, because we always employ finite sets of premises of finite length. This is not at odds with (**CnA1**), which admits that the set of sentences can be denumerably infinite, because it means that this set can simply be inductively extended; even if we admit that $\aleph_0$ represents actual infinity, it is a fairly moderate ontological presupposition. (**CnA2**) is obvious as including axioms as well as other earlier asserted assumptions among theorems. (**CnA3**) says that *Cn* acting more than once on a given set, produces nothing more. The problem of monotonicity (see (**CnA4**)) is more complicated and I restrict myself to one only remark in favour of this property, namely that it is plausible to say that if we can derive something from the empty set, it is also derivable from any other set. Let us take for granted that (**CnA1**)–(**CnA5**) are justified (I do not suggest that the proposed justification is absolute). The remainder **Cn**-axioms characterize classical logic. If they are changed—for example, by weakening the force of negation—a non-classical logic is obtained, for example, intuitionistic. The deduction theorem is, of course, very desirable. In particular, it is essential for obtaining (**DfLOG1**).

However, (**DfLOG1**) applies not only to **FOL**. Leaving aside non-classical cases, this definition is equivalent (see Surma 1981) to two other statements, namely:

(**DfLOG2**) $A \in$ **LOG** $\Leftrightarrow \neg A$ is inconsistent.
(**DfLOG3**) **LOG** is the only non-empty product of all deductive systems (theories).

In order to obtain a justification of these two definitions and their equivalence to (**DfLOG1**), (**CnA1**)–(**CnA9**) are enough, and, moreover, since, we have intuitionistic counterparts of (**CnA8**) and (**CnA9**), all three characterizations apply to intutionistic logic (I recall that classical logic is the main target here). (**DfLOG2**) and (**DfLOG1**) define the properties which that we expect to be possessed by any reasonable logic (paraconsistency is to be separately discussed at this point, but I leave this issue aside). We agree that negations of logical principles are contradictory and that logic is the common part of all theories, even mutually, inconsistent. Additionally, (**DfLOG3**) entails that logical laws are derivable from arbitrary premises. Thus, we immediately obtain the assertions:

(32) $A \in Cn\varnothing \Leftrightarrow A \in Cn\mathbf{X}$, for any $\mathbf{X}$,
(33) $\mathbf{LOG} = Cn\varnothing = Cn\mathbf{X}$, for any $\mathbf{X}$.

Yet one may suggest that the above explanations seem to play with **FOL** and **LOG** in a way—sometimes regarding them as interchangeable, sometimes not. Moreover, every formal system can be defined as $Cn\varnothing$, if the axioms for Cn are modified. Let **Th** be a theory axiomatized by a set **Ax** of axioms, and let the symbol **CA** refers to the conjunction of the axioms of **Th**. Assume further that $A \in Cn\mathbf{Ax}$. By the deduction theorem we have $(\mathbf{CA} \Rightarrow A) \in Cn\varnothing$. This is all right. However, if we add the formula (a) $\mathbf{CA} \in Cn\varnothing$ as a new axiom for Cn, we obtain that $A \in Cn\varnothing$. On the syntactic level, nothing precludes such moves. In fact, the axiom (**CnA12**) is of this kind. It was added because there are reasons for considering identity as a logical concept. However, it is difficult to agree that the axioms of type (a) are always sound as ingredients of logic. In most cases they are not. These remarks suggest that it is significant to have another account of logic that would be independent of the path that proceeds *via Cn*. Semantics motivates

(**DfLOG4**) $A \in \mathbf{LOG} \Leftrightarrow$ for every model $\mathbf{M}$, A is true in $\mathbf{M}$.

This definition describes logic as consisting of laws that are true in every model (domain, possible world, interpretation of extralogical vocabulary, etc).

We can now return to the universality property of logic. I distinguished four ways of understanding this property. To repeat: (**Un1**) logic is universal, because it is universally applicable; (**Un2**) logic is universal, because it is topic-neutral; (**Un3**) logic is universal, because its principles are universally valid; (**Un4**) logic is universal, because it has great expressive power. I also suggested that (**Un1**)–(**Un3**) are mutually equivalent. I shall now proceed to a precise formulation of these intuitions. If **LOG** is a part of every theory, it means that it is universally applicable, that is, in every concrete field. Exactly the same follows from (**DfLOG4**), because logic, as true in every model, is applicable in every concrete deductive inference. Further, since **LOG** belongs to every theory **Th** independently of **Th**-content, **LOG** is true in every model, does not depend on specific assumptions, and it is also topic-neutral. Thus, starting from (**Un1**) or (**Un2**) or (**Un3**), we intuitively (by an informal reasoning) obtain the other points. Formally speaking, (**Un1**)–(**Un3**) are equivalent, at least, if we can accept that (**DfLOG1**) and (**DfLOG4**) are equivalent as well, that is,

(34) $A \in Cn\varnothing \Leftrightarrow$ for every model $\mathbf{M}$, A is true in $\mathbf{M}$.

The justification of (34) follows immediately from the completeness theorem. However, recall that it has two versions: strong (**SV**) and weak (**WV**). The latter is more attractive here, because it pertains to strictly logical systems (consisting exclusively of tautologies). On the other hand, (**SV**) leads to the following definition of logic

(**DfLOG5**) **LOG** = <**L**, *Cn*>,

where **L** is an arbitrary first-order language. According to (34), *Cn* in its right side should be replaced by $\models$ on the semantic level. There is, of course, nothing wrong with looking at logic as an arbitrary first-order language together with a consequence operation, but that does not deal directly with the universality of logic. Assume that a **LOG** satisfies (**SV**), (**CnA5**) and (**CnA7**). Consequently, every derivation in **LOG** is reducible to a derivation from a finite set of premises, and the right side of (**SV**) can be replaced by $\mathbf{CX} \vdash A$, where $\mathbf{X} \in \mathbf{FIN}$. By (**CnA7**), that is, (**DT**), we obtain $\varnothing \vdash \mathbf{CX} \Rightarrow A$ and, further, by (**WV**)—that the implication (a) $\mathbf{CX} \Rightarrow A$ is universally valid. Therefore, (a) is a tautology. In fact, (**SV**) says that a derivation represented by $\mathbf{CX} \vdash A$ proceeds *via* a rule of logic, which is represented by a logical theorem (a). The universality property of (a) is directly established by (**WV**). Although (**WV**) is obtainable from (**SV**), the former still says something non-trivial, particularly about the universality of logic. It seems that (**WV**) is philosophically much more important for logic conceived as a collection of tautologies. Having justified (**DfLOG4**), it is easy to show that (a)–(c) express the same property. (**DfLOG1**) says that logic is independent of any specific assumptions. It is formally displayed just by the first definition of logic) and its corollary, which says that logic is a part of every theory. (**DfLOG4**) indicates that logical laws are universally valid and topic-neutral. Now (**VW**) establishes that (**Un1**)–(**Un3**) are equivalent, not only by convention, but due to a firm metalogical result, that is, by the weak completeness theorem. In fact, (**WV**) is not the only metalogical result that displays the universality property. An additional hint comes from (31), because if logic does not distinguish any extralogical content, it is just universal—in particular, neutral with respect to specific topics (domains of a special interest).

I am inclined to say that (**SV**) is about *logica utens*, but (**WV**) about *logica docens*. However, I will argue that both notions are in a sense equivalent. The former consists of rules of inferences, the latter of theorems. Let $\mathbf{LOG}^{\mathbf{R}}$ consists of a collection of rules and $\mathbf{LOG}^{\mathbf{T}}$ covers a class of theorems. Assume that $\mathbf{R} = <\{A_1, \ldots, A_n\}, A>$ is a rule of inference with premises $A_1, \ldots, A_n$ and the conclusion A. The deduction theorem and (**SV**) justifies

$$(35) \quad <\{A_1, \ldots, A_n\}, A> \ \in \mathbf{LOG}^{\mathbf{R}} \Leftrightarrow (A_1 \Rightarrow (\ldots \Rightarrow (A_n \Rightarrow A)\ldots)) \in \mathbf{LOG}^{\mathbf{T}}.$$

This establishes the parity of $\mathbf{LOG}^{\mathbf{T}}$ and $\mathbf{LOG}^{\mathbf{R}}$ and thereby also the parity of *logica utens* and *logica docens*. It means that the definitions (**DfLOG1**)–(**DfLOG4**) can be applied to logical theorems as well as to logical rules.

The normativity of logic (how should we think in order to be logical?), as related to (**DfLOG4**), has an interesting feature, which is related to Frege's point (see Chap. 3, Sect. 3.6; I slightly change the symbolism here) that logic tells us how to think in order to attain truth. Since logic does not favour any possible world (model), every world is logically accessible from any other. The standard definition of obligation tells us that **O**A (it is obligatory that A) is true in our world $\mathbf{W}^*$ if and only if A is true in all possible worlds accessible from **W*** (see Chap. 4(52)). If A is a tautology, it is true in all worlds, including **W***. Thus **O*t*** (where the symbol ***t*** denotes an arbitrary tautology **O**—is read 'it is obligatory that') is true in **W*** (in any other world as well). Thus, tautologies generate the realm of logical oughtness (we do not worry about the ontological status of this realm). Further, the relation of logical accessibility is reflexive. It means that **O*t*** implies ***t***. We also have the reverse dependence. Briefly, 'ought' and 'is' are not distinguishable in logic. This can be interpreted as an exception to the Hume thesis (see Chap. 4, Sect. 4.8) that ought is logically separate from is, but it appears to be the only exception. I suggest that it is a proper interpretation of Frege's idea that logic is normative (see Woleński 2016a).

Frege argued that if A is true, we should assert A. Hence, since tautologies are true, we should assert them unconditionally. Logic in itself does not force anybody to assert it, but when it comes to the cognitive game, the situation changes—because the obligation to assert something shows up. On the other hand, truth and assertion are not the same (Chap. 4, Sect. 4.8), because if A is asserted, it does not need to be true, unless we are dealing with tautologies. Therefore, Frege's opinion that logic is normative, because it says how we must think in order to attain truth, has to be somehow corrected. As matter of fact, attaining truth is concerned, the normativity of logic is restricted to inferences. Assume that Cn (as a closure operator) closes assertion. This simply means that if A is asserted (the grounds of extralogical assertions are not relevant here), $B \in Cn\{A\}$, then B is asserted. In order to be more realistic, we can add that it is known to the inferring person that $B \in Cn\{A\}$. Assume that A is asserted, it is known that $B \in Cn\{A\}$ and B is not asserted. By the deduction theorem, we obtain $(A \Rightarrow B) \in Cn\varnothing$. Thus, the formula $(A \Rightarrow B)$ is a tautology. Applying the principle of the assertion of tautologies, we obtain that the formula $(A \Rightarrow B)$ is asserted $(\mathrm{As}((A \Rightarrow B))$. Since assertion is distributive over implication, we get $\mathrm{As}(A) \Rightarrow \mathrm{As}(B)$. However, if B is not asserted, A is also not asserted, contrary to the first assumption. An easy argument shows that if A is asserted conditionally as the conclusion of a correct inference with asserted premises, it ought to be asserted as well.

FOL consists, as *logica docens*, of tautologies and, since it satisfies (**WV**), has the universality property in the sense of (**Un1**)–(**Un3**). There is something more to be said about **FOL** in the light of (**WV**). Let me note at first that (**DfLOG5**) is also applicable in this case, but with the proviso that **L** has a purely logical vocabulary, that is, individual variables, propositional connectives, quantifiers, identity, and predicate letters understood as non-specified parameters. We assume that a logical theorem is a formula consisting of the above building blocks and derivable from the empty set of premises. Clearly, logic as a system of theorems is never systematized by listing all logical tautologies; that would even be impossible, since there are infinitely many logical truths. Hence, logic is codified by a suitable axiomatic system. In particular, **FOL** has a complete axiomatization. Assume that $\mathbf{Ax}^{\mathbf{FOL}}$ is a set of axioms for first-order logic, that is, $\mathbf{FOL} = Cn\mathbf{AX}^{\mathbf{FOL}} = Cn\varnothing$. This means that the *logica docens* is generated from the axioms. Hence, the rules leading from axioms to theorems must preserve tautologicity, although the rules associated with logic understood as <**L**, $\vdash$ > preserve extralogical truth as well. Although a tautologicity-preserving rule is also truth-preserving, the inverse connection does not hold. Rules for <**L**, $\vdash$ > act as infallible, that is, transmit truth from premises of inferences to their conclusions, but rules for **LOG** = $Cn\varnothing$ transmit tautologicity. Although truth-preservations is sufficient for logic as <**L**, $\vdash$ >, where **L** is arbitrary (but first-order), the difference touched on in this section is essential, at least from the philosophical point of view, because it points out a certain important feature of universality as a property of *logica docens*.

The foregoing discussion suggests the first-order thesis, that is, a philosophical solution to how logic is to be understood:

(**FOT**) **FOL** is *the* logic.

his thesis is strongly contested at present. Before I pass to details, I would like to consider the following characterization of logic (Westerståhl 1976, pp. 7–8; I omit the issue of extensionality, because it is not relevant in the present context):

> (1) The study of logic is the study of a certain type of concepts, most important of which are the concept of logical consequence and logical truth. [...]. Put differently, it is the study of theories or instrument of *deduction*. [...].
>
> (2) Logical truth is truth due (only) to logical form. [...].
>
> (3) Truth is a relation between sentences on the one hand and the structures on the other [...]. [...].
>
> (4) In logic there are not privileged objects.

Further (p. 16), Dag Westerståhl defines logic as an ordered pair <$\mathbf{S_L}$, $\models$>, where $\mathbf{S_L}$ the class of sentences of a language **L**, and $\models$ is the truth relation (I omit additional constraints concerning morphisms between structures that interpret **L**); of course, the pair <$\mathbf{S_L}$, $\models$> is equivalent to the pair <**L**, $\models$>, since **L** is understood in this book as a set of sentences.

It is clear that Westerståhl combines various accounts of universality. Point (4) gives a version of the thesis that logic is topic-neutral; points (1)–(2) are familiar from the previous remarks. The view expressed in (3) stresses the semantic nature of the concept of truth. Now, Westerståhl's definition of logic is related (I am not sure whether consciously or not) to the mentioned idea of *characteristica universalis* (*logica magna*) in its moderate version. In fact, Westarståhl develops the idea of abstract logic understood as a collection of formal schemata constructed in order to investigate various mathematical structures. Thus, we obtain a new characterization of an old idea, which consists in the semantic explication of the nature of *logica magna* as language *cum* the satisfaction relation. Although this explanation of the nature of logic is interesting in itself, it does not contribute very much to the problem of how (**Un1**)–(**Un3**) and (**Un4**) are related. Generally speaking, we have two objects: (I) <**L**, $\models$> and (II) <**L**, $\vdash$> The question concerning the relation of (I) and (II) is good, because, as we shall see, the answer to it opens the way to a promising account of the universality property of logic. **FOT**, as I already demonstrated, claims that both accounts of logic are equivalent. It means, the rejection of this thesis means that the equivalence in question does not hold generally. This is Shapiro's view (see Shapiro 1991, Shapiro 1996a, p. XV) that first-order logic is a calculus only. However, **LOG** = $Cn\varnothing$ is also a language with the satisfaction relation, because the completeness theorem also allows to see it as <**L**, $\models$>.

Let me stress here that the distinction between (I) and (II) does not mean the same as that between *logica utens* and *logica docens*. On the contrary, the people who reject **FOT** insist that we need a powerful expressive scheme just because logic as a codification of deductive means (usually identified with **FOL**) has a very limited application. A message of this kind is clearly indicated by the following words (Barwise 1985, pp. 5–6, p. 23):

> As logicians we do our subject a disservice by convincing others that logic is first-order logic and then convincing them that almost none of the concepts of modern mathematics can really be captured in first-order logic. Paging through any modern mathematics book, one comes across concept after concept that cannot be expressed in first-order logic. Concepts from set theory (like *infinite set*, *countable set*), from analysis (like *set of measure* 0 or *having the Baire property*), from topology (like *open set* and *continuous function*), and from probability theory (like *random variable* and *having probability greater than some real number r*), are central notions in mathematics which, on the mathematician-in-the-street view, have their own logic. Yet none of them fit within the domain of first-order logic. In some cases the basic presuppositions of first-order logic about the kinds of mathematical structures one is studying are inappropriate (as the examples from topology or analysis show). In other cases, the structures dealt with are of the sort studied in first-order logic, but the concepts themselves cannot be defined in terms of the "logical constants". […]. Extended model theory adds a new dimension and new tools to the study of the logic of mathematics. The first-order thesis, by contrast, confuses the subject matter of logic with one of its tools. First-order logic is just an artificial language constructed to help investigate logic, much as the telescope is a tool constructed to help study heavenly bodies. From the perspective of the mathematician in the street, the first-order thesis is like the claim that astronomy is the study of the telescope. Extended model theory attempts to take the experience gained in first-order model theory and apply it in ever broader contexts, by allowing richer structures and richer ways of building expressions. It attempts to build languages similar to the first-order predicate calculus to study concepts that are banned from logic by the first-order thesis. […] Mathematicians often lose patience with logic simply because so many notions from mathematics lie outside the scope of first-order logic, and they have been told that that *is* logic. The study of model-theoretic logics should change that, by getting at the logic of the concepts mathematicians actually use, by finding applications, and by the isolation of still new concepts that enrich mathematics and logic. […] one thing is certain. There is no going back to the view that logic is first-order logic.

Barwise's rejection of **FOT** is explicit and radical. His main argument appeals to the very poor applicability of **FOL** in mathematics. His arguments are pragmatic, because they point out that **FOL** is not suitable for defining and analyzing mathematical structures and mathematical concepts. In particular, extended model theory (other labels: abstract model theory, abstract logic) is of the utmost significance for mathematics, because it increases considerably the expressive power of logic. In fact, Barwise does not claim that **FOL** is to be rejected, but argues that it is not sufficient "from the perspective of the mathematician in the street" and must be enriched by devices offered by extended model theory. This leads to so-called abstract logic. "To put **FOL** in its right place" can serve as a concise summary of Barwise's position toward **FOT** and first-order logic. He is right about the limitations of first-order logic. Although $\mathbf{L_{FOL}}$ is much more powerful than $\mathbf{L_{PC}}$, its expressive devices doe not suffice for mathematics, for example, to define the concept of finitude. However, a simple retort to Barwise is to point out that (**Un1**)–(**Un3**) are at odds with (**Un4**). To repeat, metaphorically speaking, the universality-property as defined by (**Un1**)–(**Un3**) is inversely proportional to universality as a measure of content. Consequently, if someone selects the universality property in the sense of (**Un4**) as a guide for one's philosophical orientation, on must abandon **FOL** in favour of other systems—for example, second-order logic or infinitary logic. On the other hand, the defenders of **FOT** argue that **FOL** has various elegant and nice properties. In particular, it is semantically complete and has an effective (recursive) proof-procedure—contrary to second-order logic.

As in the case of other philosophical issues, we are faces with a choice between **FOT** and its negation. I choose the former. Thus, let us continue discussion from this point of view. Is (34) a sufficient and necessary condition as criterion of logic? Certainly, it is a necessary condition. As such it excludes second-order logic, because its completeness theorem does not treat all models *al pari*. More specifically, second order logic with full models is incomplete, but it becomes complete if its models are in some way stratified. However, second-order logic in the later case is equivalent to many-sorted **FOL**. Thus, second-order logic with standard semantics (no model is distinctive) is not universal owing to its incompleteness, but it is also not universal when non-standard (Henkin) semantics is admitted for the stratification of models (it should be considered as a first-order extralogical theory). Boolos (see Boolos 1975, p. 77, page-reference to the reprint) tries to overcome this argument. He says:

> I know of no perfectly effective reply to this view [that logic is topic-neutral – J. W.]. But, in the first place, one should perhaps be suspicious of the identification of subject matter and range. (Is elementary arithmetic really not *about* addition, but only *about* numbers?) And then it might be said that logic is not so "topic-neutral" as it is often made out to be: it can easily be said to be about the notions of negation, conjunction, identity, and the notions expressed by "all" and "some", among others (even though these notions are almost never quantified over). In the second place, unlike *planet* or *field*, the notions as of *set*, *class*, *property*, *concept*, and *relation*, etc. *have* often been considered to be distinctively logical notions, probably for some such very simple reason that anything whatsoever may belong to a set, have a property, or bear a relation. That some set- or relation-existence assertions are counted as logical truths in second or higher-order systems does not, it seems to me, suffice to disqualify them as systems of logic, as a system would be disqualified if it classified as a truth of logic the existence of a planet with at least two satellites.

I must remark that logic is not about logical concepts. They are studied in metalogic. Take the notion of conjunction. As it was shown in Sect. 5.3, it can be construed as a function from **L** to the set **{1, 0}**. If $A, B \in \mathbf{L}$, then $\mathbf{v}(A \wedge B) = \mathbf{1}$ if and only if $\mathbf{v}(A) = \langle \mathbf{S_L}, \models\rangle (B) = \mathbf{1}$; otherwise, $\mathbf{v}(A \wedge B) = \mathbf{0}$. Yet no theorem of propositional logic asserts that conjunction behaves in this way. We should rather say that the formula $A \wedge B \Rightarrow B$ becomes universally valid according to the above definition of conjunction. The second argument also has a very weak force, because only its content is subject to the controversy in question. The argument (see Corcoran 2001) pointing out the indispensable role of second-order sentences (for example, 'true sentences logically imply true sentences') in elaborating properties of first-order tautologies seems to confuse logic and metalogic.

On the other hand, (34) does not provide a necessary condition, because there are logics other than **FOL** which are semantically complete, for example, some infinitary logics or logics with infinitary rules—say the ω-rule (it is a rule having infinitely many premises). However, if we say that (**CnA5**) is a natural property of logic, then only **FOL** remains. Thus, *the* logic, on the proposed views, has two marks, namely the universality property and the finitary character of the inference rule. Note, however, that cancelling (**CnA5**) as a source of logical properties still gives a definition of logic that ascribes the universality property in the considered sense to some other systems than **FOL**. But if some generalized quantifiers (for instance, 'there exist countably many' or 'there exists uncountably many') are added, (31) does not hold and the universality property is violated, because the resulting logics lose topic-neutrality. This suggests, as opposed to many contemporary proposals (see Westerståhl 1976), that generalized quantifiers, as favouring some cardinalities of sets, are not logical constants–contrary to the usual quantifiers, that is, 'for every' and 'there is'. Perhaps another argument (due to Alexander M. Levin and pointed out to me by Valentin Shehtman) casts additional light on this point. Logic should take into account the absolute properties. However, due to the (**LS**), the notion of cardinality is not absolute. Hence, any theory that distinguishes various cardinalities is not a logic in the proper sense. As far as the quantifiers, 'for every' and 'there is' (in particular, the former) are concerned, they appear as the only purely logical ones, contrary to numerical quantifiers, for example.

One can ask what (30) tells us about the universality property. First of all, the Lindström theorem concerns rather *logica utens* (in the semantic version), that is, $<\mathbf{L}, \models>$, than *logica docens*. Secondly, (30) addresses to the expressive power of logic rather than its universality property. It is of course very interesting that completeness and compactness act to the same effect when they occur together with the Löwenheim property. If logic is understood as $Cn\varnothing$, its compactness is a trivial property. Applying it to the universality property, we obtain that a set of sentences is universally valid if and only if every finite subset of it is universally valid, but this is nothing surprising (see also Appendix). The Löwenheim property displays an aspect of universality, namely that **FOL** does not distinguish between models with different cardinalities (however, see below). This feature of **FOL** is also not surprising, because it treats all models *al pari* modulo the satisfaction relation. Thus, compactness and the Löwenheim property are fairly natural from the point of view of *logica docens*, if it is identified with **FOL**. Its expressive power is indeed very poor, but certainly not null. At first, we assume that models of **FOL** are not empty. Secondly, **FOL** discriminates syntactically various constants and predicates by indexing their places and arities. If identity is present, numerical quantifiers could be added. Thirdly, denumerable cardinality is distinguished due to the Löwenheim–Skolem–Tarski theorem. It is really a very surprising fact that if something is first-order satisfiable at all, it is satisfiable in a denumerable domain, even if we explain this by the cardinality of **L**, but this fact is generated by the syntactic properties of first-order languages. These remarks confirm that the feeble expressive power of **FOL** as the set of tautologies is a cost of its universality property. Clearly, **FOL** has various limitations—those displayed by (30), in particular—but it can be considered just as *the* logic, if the universality property is taken as the measure.

The above considerations should be supplemented by the remark that our semantics is based on standard set theory. That is not without importance. One can ask how reliable set theory is as the basis of a semantics or metalogic for **FOL**. Of course, its reliability does not exceed that witnessed in other parts of ordinary mathematics. However, when using standard set theory in the semantics and metalogic of **FOL**, we need to employ only a part of the set theoretical universe (in fact, the weak second order arithmetic with the axiom of arithmetical comprehension is enough for first-order model theory; see Murawski 1999, Simpson 1999, Halbach 2011). This circumstance is related to the absoluteness of **FOL** (see Väänänen 1985, Vänäänen 2001; this second paper shows how the absoluteness of **FOL** is related to (3)). Roughly speaking, a logic (in the sense of $<\mathbf{L}, \models>$) is absolute if the truth-value of the expression $\mathbf{M} \models A$ depends on the existence of some selected sets (the existence of such sets is guaranteed by the arithmetical comprehension axiom). On the other hand, second-order logic is not absolute in this sense, because it generates problems connected with the continuum hypothesis and other independent set-theoretical statements (see Vänäänen 2001, Andréka, Madárasz, Németi 2003). If we proceed to metalogic of second-order logic, we cannot neglect the differences between various possible extensions of **ZFC**. In particular, Vänäänen argues that the 1st order **ZFC** is just as good as second-order logic. Since the latter operates with a very relative notion of set, this makes it impossible to decide on clear logical grounds which set-theoretical universe is really "good". Thus, according to Vänäänen, it is quite illusory to maintain that second-order logic gives us *the* proper characterization of the set-theoretical universe. Certainly, it is possible to appeal to other metalogical schemata or to universal algebra (see Beziau 2007; the program of universal logic modelled on universal algebra), category theory (see Goldblatt 1979; categorial logic) or abstract algebraic logic (see Font 2016)—but I do not think that that would change the situation in a radical way.

If we pass to *logica utens*, that is $<\mathbf{L}, \vdash>$ and $<\mathbf{L}, \models>$ what is natural from the perspective, *logica docens* might be seen otherwise from the point of view of applications. Although I have no ambition to introduce terminological innovations, let me temporary speak about first-order formalizations (**FOF**) of theories, instead of first-order logic. In order to display this idea in an explicit manner, let $<\mathbf{L}, \vdash>$ and $<\mathbf{L}, \models>$ be replaced by $<\mathbf{L^1}, \vdash>$ and $<\mathbf{L^1}, \models>$, where the superscripts refer to the order of language. Logic is therefore hidden in $\vdash$, and semantics in $\models$. Now, we see that one should not speak about the expressive power of $\vdash$ or $\models$, but refer this capacity to $\mathbf{L^1}$. There is no doubt that the expressive power of this language is very limited, but it is fairly independent of the properties of the consequence relation.

This limited expressive power is also responsible for the non-categoricity of **FOF**, that is, for their having non-isomorphic models (this is the consequence of (29)), contrary to second-order formalizations. In spite of the virtues of higher-order languages, the matter of whether **FOF** are good (or, how good) for mathematics and science is still a controversial issue (see Vänäänen 2001, Andréka, Madárasz, Németi 2003 for defence of **FOT**, and the quoted works of Shapiro for the opposite view) and must be omitted here. Since I am primary interested in the concept of logic, and not in **FOT** as providing a language for mathematics, I see (**Un1**) and its equivalents as a fundamental property of *the* logical. If some logicians, want to have expressively powerful languages, they have to abandon **FOF** in favour of other formalisms, for instance, second-order ones. This move leads to systems with the universality property in the sense of (**Un4**), and confirms in addition that universality in this sense is at odds with the universality property as characterized by (**Un1**)–(**Un3**) and formally displayed by the metalogical characterization of **FOL**. Moreover, observe that $<\mathbf{L}\ \models>$ is not comparable with $<\mathbf{L},\ \vdash>$ without appealing to metalogical properties. Thus, we have a kind of dialectic between the universality property as validity (and its cognates) and universality as expressive power. If a logician wishes to have both universalities without any cost, this task seems to be rather fanciful impossible and, in speaking of logic, one needs to make a commitment either the universality property, or the great expressive power.

Incidentally, the opinion that second-order theories are categorical is misleading to some extent. First of all, they are incomplete by the first Gödel theorem, because if arithmetic is consistent, its extensions obtained by adding undecidable sentences, are also consistent and have models (see Chap. 8, Sect. 8.6). However, these models are radically different, even though they can be equicardinal, which means second-order theories also have non-standard models, and this fact seems to be a derivative of the powerful expressive devices of second-order languages. Now, if it is the case, the objection which points out that first-order theories do not distinguish standard and non-standard models with respect to their cardinality is simply unfair. There is no actual doubt as to which models are standard from the first-order perspective. The mistake consists here in an unfounded belief that we have purely logical criteria for what makes any model standard in the case of second-order theories. In fact, all employed criteria of being a standard model for a given theory, for instance arithmetic, are always extralogical. Anyway, one should make a choice concerning which formalism is to be used in given investigations. As I already noted, I opt for first-order formal tools as basic from the point of view of logic. This preference is also motivated by so-called Hilbert thesis, stating that every higher-order language can be elementarized, that is, transformed into first-order one. Of course, it costs something, for example, a complication of axioms, but this eventually unpleased factor finds its compensation in nice properties of first-order schemata. To conclude, every view about what is logic, or which system should be regarded as *the* logic, is deeply rooted in a philosophical environment. My presumptions (or prejudices if someone wants) determine **FOT**.

How look at the variety of logics in the light of the definition (**DfLOG1**)? Take the dual logic for example. We can define it as $\boldsymbol{d}Cn\varnothing$. That is not surprising, because dual logic is a mirror of **FOL** via duality. Intuitionistic logic can be presented via *Cn* acceptable for the intuitionists, and is definable as the set of intuitionistic consequences of the empty set. The same can be done for many-valued logic, fuzzy logic, paraconsistent logic, non-montonic logic or relevant logic, provided that a suitable *Cn* were to be axiomatized. However, this success is limited, because criteria for distinguishing logical and extralogical ingredients in possible definitions of *Cn* are really vague. Consider, for example, modal logics in this context. Their possible-world semantics are based on various formal properties of the accessibility relation like reflexivity, symmetry, etc. However, these properties are not sources of the required universality (some semantic modal constructions have given properties, others not). Thus, they introduce some extralogical element to the logical behaviour of modalities. In fact, only the system called **K** treats all possible worlds equally (no additional constraints on the accessibility relation are imposed). Its modalities, that is, necessity and possibility, behave exactly like quantifiers. Perhaps this system represents the pure modal logic. However, and it is not surprising, **K** is of this system not sufficient in order to cover all intuitions connected with modalities, because the expressive power of this system is relatively small. This fact gives reason for considering logics of modalities (at least those that have a richer content than ones formalized by **K**) seem to rather formal extralogical systems rather than as purely logical. Clearly, every logical system is formal, but the reverse implication does not hold.

As the last paragraph shows, the question 'Which logic is the right logic?' cannot be reduced to the choice between **FOL**, second-order logic or infinitary logic. Another issue focuses on the rivalry between classical logic, its extensions and its various non-classical alternatives. The typical way of discussing this problem consists in the following question: "Can or should we replace classical logic by some other system, for instance, intuitionistic, many-valued, relevant or paraconsistent logic?" This way of stating the problem distinguishes classical logic as the system which serves as the point of reference. Thus, alternative or rival logics are identified as non-classical. There are two reasons in order to regard classical logic as having a special status. One reason is that classical logic appeared as the first stage in the development of logic. This argument is not particularly strong, because it refers to historical (genetic) and purely descriptive circumstance. The second motive is clearly evaluative in its character and consists in saying that classical logic has the most "elegant" properties or that it "best" serves for science, and mathematics, and perhaps for ordinary arguments as well. It is said, for example, that there is something wrong about abandoning the principle of excluded middle (intutionistic logic and other constructive systems), introducing more than two logical values (many-valued logic), changing the meaning of implication (relevant logic), or tolerating inconsistencies as not dangerous (paraconsistent logic). It is also often argued that some non-classical logics—say intuitionistic or many-valued logics—restrict considerably the applicability of formal logic to mathematics as it is accepted by working mathematicians. This argument is perhaps the most dramatic in the case of intuitionistic logic, because it leads to eliminating a considerable part of classical mathematics, for instance, proofs based on the axiom of choice and other non-constructive devices. Consequently, so the reported argument proceeds, only classical (bivalent) logic adequately displays the proof methods employed in ordinary mathematics as it is. And as long as the discussion on the rightness of this or that logic is conducted in descriptive language, it appeals to intuitive assessment of what is right or wrong in mathematics.

Another controversy concerns the actual role of classical logic in proving metalogical properties of particular logic systems, features such as completeness, decidability, and like. The priority of classical logic is sometimes explained by pointing out that some properties of non-classical logic are provable only classically. This is well-illustrated by the case of the completeness of intuitionistic logic: Is the completeness theorem for this logic intuitionistically provable? The answer is not quite clear, because the stock of intuitionistically or constructively admissible methods is not univocally determined, and they vary from one author to another. Anyway, most authors agrees that it is problematic whether the intuitionistic proof of the completeness theorem for intuitionistic logic is possible at all. An interesting fact is that in the case of **FOL**, **(SV)** and (26) (the Gödel–Malcev theorem) have non-constructive proofs only, but their equivalence is intuitionistically provable. This example shows that the interplay between classical and constructive aspects of metalogical (or metamathematical) statements is fairly deep and cannot (with awareness that all predications what is impossible must be taken *cum grano salis*) be solved by very general philosophical assumptions.

Finally, our main problem (what is logic and what is its scope?) is also connected with the extensions of logics. When we construct modal logics, deontic logics, epistemic logics, etc., we usually start with some basic amount of logic, propositional or predicate, classical or not. Consequently, we have modal propositional, or predicate systems that are based on (or are extensions of) classical, intutionistic, many-valued, paraconsistent, or some other basic logic. Does any given extension (roughly speaking, an extension of a logic arises when we add new concepts—say necessity—to an old ones, in such a way that all theorems of the system before extension are also theorems the new system) of a chosen basic logic preserves its location as a genuine logic, or does it produce an extralogical theory? The *a priori* answer is not clear, even when we decide that this or that basic system is *the* logic. The problem of the status of extensions of logic(s) is particularly important for so-called philosophical logic because it concerns mainly systems belonging to this area which occupies a territory between pure logic and philosophy.

5.5 Are There Degrees of Logicality?

Could we possibly distinguish various degrees of being logical? This is suggested by the fact that **FOL** has three segments (I consider this problem from the point of view of axioms for *Cn*, but I omit general axioms):

(A) propositional calculus (axioms (**CnA6**)–(**CnA9**));
(B) first-order predicate logic without identity (axioms (**CnA6**)–(**CnA11**);
(C) identity (axioms (**CnA6**)–(**CnA13**)).

As I already noted, there are properties possessed by some parts that are not attributable to others. (A) is semantically complete, Post-complete and decidable. (A) + (B) and (A) + (B) + (C) (the full FOL) are neither Post-complete nor decidable. (A) + (B) obeys the inflation and the deflation theorems, whereas these theorems do not hold for (A) + (B) + (C). If one insists that decidability is a natural property, only propositional logic (possibly plus some fragments of predicate calculus) remains as *the* logic. If one maintains that the inflation and deflation theorems introduce too much extralogical content into logic, only first-order logic without identity remains. On the other hand, the systems (A), (A) + (B) and (A) + (B) + (C) are semantically complete and, as I already suggested that is perhaps their most important logical property (they are, of course, also consistent, but this attribute has no use in determining which one is *the* logic).

Now I shall try to show that, although the set of tautologies of **FOL** is not maximally (Post) consistent, there is another meaning of maximality that can be attributed to first-order logic as determined by (A) + (B) + (C). Firstly, note that the consequences of tautologies should be tautologies too (recall that proofs inside $Cn\varnothing$ preserve tautologicity). We need to fix the semantic status of the empty set of sentences. This set is finite, but is not enough. Every finite set **X** of sentences is representable by a finite conjunction **CX**. Since $\mathbf{CX} \in Cn\mathbf{X}$, for any **X**, we have that $\mathbf{C}\varnothing \in Cn\varnothing$. The assumption that we are just working in a logic that satisfies the (**WV**), implies that the formula $\mathbf{C}\varnothing$ is a tautology. Since $\mathbf{C}\varnothing$ represents the set $\varnothing$, the latter has to be regarded as the tautological set. In other words, we have the property **TAUT** associated with the set $Cn\varnothing$ such that **TAUT**(A) if and only if for any model **M**, A is true in **M**. It is clear that:

(35) (a) $\mathbf{TAUT}(\varnothing)$;
(b) $\mathbf{TAUT}(Cn\varnothing)$ if and only if $\mathbf{TAUT}(\mathbf{X})$, for every $\mathbf{X} \in \mathbf{FIN}$ such that $\mathbf{X} \subseteq Cn\varnothing$.

Thus, **TAUT** is a property of finite character. If we identify the universality property with **TAUT**, this property is also of finite character. Moreover (the Tukey Lemma about finite properties), if any other set **Y** of sentences—true in all models and closed by Cn—satisfies (35a) and (35b), then $\mathbf{Y} = Cn\varnothing$. Although adding non-tautologies to $Cn\varnothing$ does not produce inconsistency in general, **TAUT** and the universality property are maximal in a well-defined sense. Observe that this reasoning does not go with respect to $<\mathbf{L}, \vdash>$ and (SV). We can naturally prove that all tautologies are universally valid and derivable from the empty set of premises, but the consequence relation preserves truth, not tautologicity. However, truth is not a property of the finite character (for a while I consider truth as a property). This remarkable property will be further elaborated in Chap. 8, Sect. 8.5. Manipulating constraints for Cn can change this situation, but it brings us back to the issue of what is natural in metalogic and what should be captured by logic; for instance, we can postulate that every truth is a consequence of the empty set.

What ten is a natural logical property? Decidability? Completeness? Maximality? The choice is open once again and cannot be made without a dose of conventionality. If we want to give the strongest possible account of 'being logical', the extension of this predicate should be determined by the properties of **PC**. **FOL** without identity provides a weaker solution, and $\mathbf{FOL}_=$ (= **FOL** as used in this book) offers a still weaker possibility, which, in my opinion, is the best—due to the fact that basic metalogical theorems hold for this system. Thus, I opt for a uniform application of 'being logical' without degrees of logicality, although I do not deny differences between PC, **FOL** and $\mathbf{FOL}_=$. It seems that every decision concerning what logic is, must be conventional to some extent. We can say that, even in **PC** (including non-classical systems), various metalogical commitments introduce some extralogical factors, for instance, that formulas are of a finite length, that paraconsistency is rejected or admitted, or that relevant implication is added. Consequently, the scope of logic depends somehow on such more or less informal constraints, supplemented by opinions what is natural in logic and what is not. This circumstance probably steered Tarski to his view (see Tarski 1936) that the borderline between logical and non-logical concepts is vague. The status of modalities is perhaps a very good example here. According to the view adopted in this book, modal logics, perhaps except the system **K**, are rather formal theories of modalities than strictly logical theories. The same applies even more strongly modal logic in an extended sense, for example, formalizing deontic, temporal, or epistemic concepts. Of course, it would be pointless to contest the contemporary usages of the noun 'logic'. On the other hand, pointing out properties that deserve to be called logical is not without rationale, even if the ultimate criterion of logicality appears utopian.

(**DG16**) I omitted here Tarski's idea of logical concepts as invariants under the class of all one-to-one transformations (see Tarski 1986), because it is applied to logic considered rather from the point of view of the theory of logical types (see Chap. 6, Sect. 6.3). This is a perspective discussing the essence of logic different from than focusing on **FOT**. In order to avoid possible misunderstanding, let me add that Tarski himself contributed essentially to the development of **FOL**.►

(**DG17**) Since discussion of the question concerning the nature of logic involves philosophy, it is difficult to expect that final solutions might be achieved even with help of metalogic. If we admit that various logics have the universal property, we must take into account that different kinds of universality can be distinguished. The next question that arises is how to compare the different meanings of universality, and which of these is basic. Thus, *res ad principiam venit* as usually happens when philosophical issues are considered. I do however think that a serious philosophical lesson can be derived from the discussion in this section. It consists of some hints on how regard logical analysis. Although it is very easy to say that we perform logical analysis, or analysis of logical concepts, or analysis of concepts via logical tools, such qualifications depend on logic is understood.►

(**DG18**) Problems discussed in this chapter have an importance for the analysis of **STT**. One of my principal reasons to discuss the nature of logic for a somewhat extensive (perhaps even excessive) way is to provide a broader perspective for considerations in the next chapters. In particular, we have a question of how logic is relevant for defining the concept of truth. If we consider **PC** truth is represented by **1** without further comments. Perhaps the following words can be regarded as symptomatic for many logicians (Lyndon 1966, p. 13):

> We want to think of an interpretation ϕ as attaching to each formula p some assertion about the structure A, which either holds or fails in A. The easiest way out is to take ϕp to be simply the value, truth or falsehood, of this assertion. Since we need not, and would rather not, explain here what is meant by truth and falsehood, we choose instead two neutral objects, the numbers 1 and 0, to serve, respectively, instead of truth and falsehood.

Anyway, neither **PC** nor **FOL** have resources for defining **1** (and **0** as well). It is possible in stronger systems, like Leśniewski's protothetic (see Chap. 3, Sect. 3.7 (F)), but, as I earlier noted, related definitions concern logical truth, not truth *simpliciter*. Inspecting (**SDfVER**) immediately shows that it is based on an extralogical condition that every object in **U** has its name. This circumstance seems to suggest that the concept of truth is not purely logical. The difference between tautologicity-preserving and truth-preserving can be regarded as an additional justification of this fact. This conclusion is important because many authors, Frege in particular (see Chap. 3, Sect. 3.6), maintain that logic is the science of truth (this statements concerns rather philosophers than mathematical logicians). However, predicates 'is a logical tautology' and 'is true' refer to different properties (although tautologicity is a special kind of truth), and the former is "more" logical than the latter. Since **STT** is frequently qualified as a logical (metalogical) theory (construction) of truth, it is important to see this qualification in a precise sense.►

(**DG18**) This chapter uses some material previously published in Woleński 2004, Woleński 2016a.►

Appendix A: Historical Note About Metalogic and Metamathematics

Alexander of Aphrodisias, an Aristotelian scholar ordered the works of the Stagirite in such a way that those of Aristotle's works devoted to first philosophy (*prote filosofia*) were placed just after the book *Physics*. Thus, the word 'metaphysics' (more precisely, its Greek counterpart) arose as a composition of 'meta' (after) and 'physics', and originally meant the same as the phrase 'after physics'. This origin of the word 'metaphysics' has no particular substantial import and indicates its role in, so to speak, librarians work. However, some historians of ancient philosophy suggest that our word was intentionally introduced to point out considerations of a certain kind, namely reflections about nature and its theory. At any rate, this more substantial use of the word 'metaphysics' quickly became the official. Today, metaphysics is considered as the theory of being, and is very often identified with ontology; eventually, ontology is understood as general metaphysics, that is, the study of the most general properties of being *qua* being.

The use of words beginning with the prefix 'meta' became quite popular in the 20th century. One can cite 'metatheory', 'metascience', 'metaethics', 'metamathematics' or 'metalogic' as examples. Their intended meaning consists in pointing out considerations about the fields indicated following the prefix 'meta'. The word 'metaphysics' would be a very good label for methodology of physics, but this use of it is excluded by the historical circumstances mentioned above—independently of whether they occurred accidentally or not. The label 'metatheory' denotes, or perhaps suggests, a theory of theories. Metascientific studies in the 20th century employed the term 'metatheory' to refer to investigations of theories in a variety of disciplines, for example, logic, sociology, psychology, history, etc.—some people claim that these investigations constitute a separate field, namely science of science. The philosophers of the Vienna Circle, who made metatheoretical studies of science the main concern of their philosophy, restricted metatheory to the logic of science modelled on developments in the foundations of mathematics. More specifically, the logic of science was intended to play a role similar to metamathematics in Hilberts's understanding that is, it was just projected as a formal analysis of scientific theories understood as well-defined linguistic entities.

The word 'metamathematics' was also used before Hilbert, but with a different meaning. In the early 19th century, mathematicians, like Gauss, spoke about metamathematics in an explicitly pejorative sense. It was for them a speculative way of looking at mathematics—something like the metaphysics of mathematics. A negative attitude toward metaphysics was inherited at that time from Kant and early German positivists. The only one serious use of 'metamathematics' was restricted to so-called metageometry. This was due to the fact that the invention of various geometries in the 19th century stimulated comparative studies. For example, investigations were undertaken of particular axiomatizations, their mutual relations, models of various geometrical systems, and attempts to prove their consistency. In this context, the word 'metageometry' referred to a well-established domain of formal studies. Presently, the prefix 'meta' means two different things. First, it indicates that metatheoretical considerations appear after (in the genetic sense) the theories that comprise the subject-matter of such studies have been formulated. Secondly, this prefix suggests that every metatheory is somehow above a theory which is investigated. It is important to see that 'above' does not function as an evaluation, but only indicates the fact that metatheories operate on a different level than theories do. A simple mark of this fact is that theories are formulated in an object–language, whereas metatheories are expressed in a related metalanguage.

Since metalogic is a part of metamathematics, it is useful to say a few words about the latter. It is probably not accidental that Hilbert passed to metamathematics through his famous study of geometry and its axiomatic foundations. Following metageometry, Hilbert projected metamathematics as a rigorous study of mathematical theories by mathematical methods. Moreover, the Hilbertian metamathematics, due to his views in the philosophy of mathematics (formalism) was restricted to finitary methods. If we reject this limitation, metamathematics can be described as the study of mathematical systems by all mathematical methods; they cover those that are admitted in ordinary mathematics, including infinitistic or infinitary; the latter freely employ, for instance, the axiom of choice or transfinite induction. However, this description is still too narrow. Hilbert's position in metamathematics can be described as follows: only syntactic or combinatorial methods are admissible in metatheoretical studies. When the Hilbertians proved theorems with semantic content about formal systems, they used semantic concepts, like validity or truth, in informal sense rather than as rigorously defined notions. Due to Tarski's works, semantics became a rigorous mathematical field and entered the domain of metamathematics (see Feferman 2004). It is perhaps interesting that the borderline between syntax and semantics corresponds to some extent to the frontier between finitary and infinitary methods. I say "to some extent" because we also have formal systems with infinitely long formulas (infinitary logic). It is clear that the syntax of infinitary logics must be investigated by methods that exceed beyond finitary tools. It was also not accidental that systematic formal semantics (model theory), which requires infinitistic methods, appeared in works of Tarski, who—due to the scientific ideology of the Polish mathematical school—did not accept the view (the finitary dogma, so to speak) that only finite methods are admissible in metamathematics (the finitary dogma, so to speak). Today, metamathematics can be divided into three wide areas: proof theory, recursion theory and model theory. Roughly speaking, the first is an extension of Hilbert's position because the above-mentioned finitistic restriction is rejected. Recursion theory is closely related to the decision problem, that is, the problem of the existence of combinatorial (algorithmic) procedures that provide methods for deciding whether a given formula is or not is a theorem. Finally, model theory, studies relations between formal systems and the structures which are their interpretations and realizations; this part of metamathematics has many affinities with universal algebra.

Metalogic is understood here as that part of metamathematics which is restricted to logical systems, and refers to studies of logical systems by mathematical methods. Of course, the scope of metalogic depends on the range of the concept of logic. For example, if **FOT** is accepted, the metalogic proper to this assumption should be restricted to the metatheory of **FOL**. The word 'metalogic' also appeared in the 19th century, although its roots go back to the Middle Ages (*Metalogicus* of John of Salisbury). Philosophers, mainly Neokantians, understood metalogic as concerned with general considerations about logic, its nature, scope, relations to other fields, etc. The term 'metalogic' in its modern sense, that is, referring to mathematical studies on logic, was used for the first time in Poland as a label for the metamathematics of the propositional calculus. Thus, metalogic is metamathematics restricted to logic, and it covers proof theory, investigations concerning the decidability problem, and model theory with respect to logic.

(**DG19**) Let me mention as curiosities from the present point of view that Ernest Troeltsch, an eminent German historian, used the term 'metalogic'as referring to methods of concrete historical investigations, and Walter Harburger—as equivalent to 'logic of music'.►

References

Agazzi, E. (Ed.) (1981). *Modern Logic*. Dordrecht: Reidel.

Anderson, C. A., Zeleny, M. (Eds.) (2001). *Logic, Meaning and Computation: Essays in Memory of Alonzo Church*. Dordrecht: Kluwer.

Andréka, H., Madárasz, J., Németi, I. (2003). Why first-order logic? (unpublished).

Barwise, J. (1985). Model-theoretic logics: Background and aims. In Barwise, Feferman (1985), 3–23.

Barwise, J., Feferman, S. (Eds.) (1985). *Model-Theoretic Logics*. Berlin: Springer.

Beall, J. C., Restall, G. (2006). *Logical Pluralism*. Oxford: Clarendon Press.

Benmakhlouf, A. (Ed.). (2004). *Sémantique et épistemologie*. Casablanca: Editions Le Fennec.

Beziau, J.-Y. (Ed.) (2007). *Logica Universalis. Toward a General Theory of Logic*. Basel: Birkhäuser.

Boolos, G. (1975). On second-order logic. *Journal of Philosophy*, 72, 509–527; repr. in Shapiro (1996), 70–87.

Borkowski, L. (1991). Characterization of quantifiers in the axiomatic theory of consequence. In Borkowski, Stępień (1991), 37–39.

Church, A. (1956). *Introduction to Mathematical Logic*. Princeton: Princeton University Press.

Corcoran, J. (2001). Second-order logic. In Anderson, Zeleny (2001), 61–75.

Feferman, S. (2004). Tarski's conceptual analysis of semantical notions. In Benmakhlouf (2004), 79–108.

Flum, J. (1985). Characterizing logic. In Barwise, Feferman (1985), 77–120.

Font, J. M. (2016). *Abstract Algebraic Logic. An Introductory Textbook*. London: College Publications.

Frege, G. (1891). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100; 25–50; Eng. tr . (by M. Black) in Frege (1984), 155–177.

Frege, G. (1892). Über Begriff und Gegenstand. *Vierteljahrsschrift für Wissenschaftliche Philosophie*, 16, 192–205; Eng. tr. (by P. Geach), in Frege (1984), 182–194.

Frege, G. (1984). *Collected Papers on Mathematics, Logic and Philosophy*. Oxford: Blackwell.

Gödel, K. (1944). Russell's mathematical logic. In Schillp (1944), 125–153; repr. in Gödel (1989), 119–141.

Gödel, K. (1989). Collected papers (v. II). Oxford: Oxford University Press.

Goldblatt, R. (1979). *Topoi. The Categorial Analysis of Logic*. Amsterdam: North-Holland.

Haack, S. (1978). *Philosophy of Logics*. Cambridge: Cambridge University Press.

Halbach, V. (2011). *Axiomatic Theories of Truth*. Cambridge: Cambridge University Press.

Hendricks, V., Neuhaus, F., Pedersen, S. A., Scheffler, U., Wansing, H. (Eds.) (2004). First-Order Logic Revisited. Berlin: Logos.

Hintikka, J. (Ed.) (1995). *From Dedekind to Gödel*. Dordrecht: Kluwer

Hintikka, J. (1996). *The Principles of Mathematics Revisited*. Cambridge: Cambridge University Press.

Hintikka, J. (1998). *Language, Truth and Logic in Mathematics*. Dordrecht: Kluwer.

Hodges, W. (2004). Which languages have Tarski truth-definition. In Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E. Romanowska, A, Woleński, J. (2004), 77–92.

Lyndon, R. C. (1966). *Notes on Logic*. New York: Van Nostrand.

Mann, A. L., Sandu, G., Sevenster, M. (2011). *Independence-Friendly Logic: A Game-Theoretic Approach*. Cambridge: Cambridge University Press.

Manzano, M. (1999). *Model Theory*. Oxford: Clarendon Press.

Murawski, R. (1999). *Recursive Functions and Metamathematics. Problems of Completeness and Decidability*. Dordrecht: Kluwer.

Parsons, D. (2016). *Theories of Intensionality. A Critical Survey*. Dordrecht: Springer.

Pogorzelski, W. A. (1994). *Notions and Theorems of Elementary Formal Logic*. Białystok: Białystok University Press.

Quine, W. v. O. (1970). *Philosophy of Logic*. Englewood Cliffs: Prentice Hall.

Schillp, P. A. (Ed.) (1944). *The Philosophy of Bertrand Russell*. La Salle: Open Court

Shapiro, S. (1991). *Foundations without Foundationalism*. Oxford: Oxford University Press.

Shapiro, S. (Ed.) (1996). *The Limits of Logic. Higher-Order Logic and the Löwenheim–Skolem Theorem*. Aldershot: Dartmouth.

Shapiro, S. (1996a). Introduction. In Shapiro (1996), XI–XXII.

Shapiro, S. (2014). *Varieties of Logic*. Oxford: Oxford University Press.

Simpson, S. G. (1999). *Subsystems of Second-Order Arithmetic*. Berlin: Springer.

Stelmach, J., Brożek, B., Kwiatek, Ł. (Eds.) (2016). *The Normative Mind*. Kraków: Copernicus Center.

Surma, S. J. (1981). The growth of logic out of the foundational research in the foundations of mathematics. In Agazzi (1981), 15–33.

Tarski, A. (1930). Über einige fundamentale Begriffe der Metamathematik. *Comptes Rendus des séances de la Société des Sciences et de Lettres de Varsovie*, 23, 22–39; Eng. tr. in Tarski (1956), 30–37.

Tarski, A. (1933). Pojęcie prawdy w językach nauk dedukcyjnych (The concept of truth in languages of deductive sciences). Warszawa: Towarzystwo Naukowe Warszawskie; Germ. tr. (with additions) as Tarski (1935).

Tarski, A. (1935). *Der Wahrheitsbegriff in den formalisierten Sprachen. Studia Philosophica*, 1, 261–405; repr. in Tarski (1986), v. 2, 51–198; Engl. tr. (*The Concept of Truth in Formalized Languages*) in Tarski (1956), 152–278 (page-references Tarski (1933)).

Tarski, A. (1936). Über den Begriff der logischen Folgerung. In *Actes du Congès International de Philosophie Scientifique*, v. 7, 1–11. Paris: Herman; Eng. tr. in Tarski (1956), 409–420.

Tarski, A. (1956). *Logic, Semantics, Metamathematics. Papers of 1923 to 1938* (tr. by J. H. Woodger). Oxford: Clarendon Press; 2nd ed., Indianapolis: Hackett Publishing Company 1983.

Tarski, A. (1986). What are logical notions? *History and Philosophy of Logic*, 7, 143–154.

Tarski, A., Vaught, R. (1957). Arithmetical extensions of relational systems. *Compositio Mathematica*, 13, 81–102; repr. in Tarski (1986), v. 4, 651–682.

Vänäänen, J. (2001). Second-order logic and the foundations of mathematics. *Bulletin of Symbolic Logic*, 7, 504–520.

Westerståhl, D. (1976). *Some Philosophical Aspects of Abstract Model Theory*. Göteborg: Institutionen för Filosofi, Göteborgs Universitet.

Woleński, J. (1995a). On Tarski's Background. In Hintikka (1995), 331–341; repr. in Woleński (1999), 126–133.

Woleński, J. (1999). *Essays in the History of Logic and Logical Philosophy*. Kraków: Jagiellonian University Press.

Woleński, J. (2004). First-order logic: (philosophical) pro and contra. In Hendricks, Neuhaus, Pedersen, Scheffer, Wansing (2004), 369–399; repr. in Woleński (2011), 61–80.

Woleński, J. (2011). *Essays on Logic and Its Applications in Philosophy*. Frankfurt a. M.: Peter Lang.

Woleński, J. (2016a). Normativity of logic. In Stelmach, Brożek, Kwiatek (2016), 169–195; repr. in Woleński (2018), 177–194.

Woleński, J. (2018). *Logic and Its Philosophy*. Frankfurt a. M.: Peter Lang.

Wójcicki, R. (1988). *Theory of Logical Calculi. Basic Theory of Consequence Operations*. Dordrecht: Kluwer.

Chapter 6
Matters of Semantics

Abstract As **STT** is a semantical theory, its presentation requires explaining what semantics is. This chapter contains a couple of historical and substantive information related to semantics, semantic antinomies, and formal languages.

6.1 Introduction

Since **STT** a semantic theory (or more carefully is interpreted in such a way), the nature and scope of semantics is of the utmost interest in the present book (Hipkiss 1995 considers definitions of semantics from various points of view, linguistic as well as philosophical; see Allan 2009 for an encyclopaedic survey of semantics, historic and substantive). It is customary to distinguish semantics sensu *largo* and semantics sensu *stricto*. The latter is usually conceived as the relation between language and what language is about. On the other semantics in the wide understanding consists of syntax, semantics sensu *stricto* and pragmatics. This tripartite division of semantics *sensu largo* is due to Charles Morris (see Morris 1938, p. 84). Carnap (see Carnap 1939, p. 146), who popularized this picture of semantic) gives the following characterization of particular subfields of semantics in the broad sense:

> If in investigations explicit reference is made to the speaker, or to put it in more general terms, tie the user of a language, then we assign it to the field of **pragmatics**. […]. If we abstract from the user of the language and analyze only the expressions and their designata, we are in the field of **semantics**. And, if finally, we abstract from the designata also and analyze only the relations between the expressions, we are in (logical) **syntax**. The whole science of language, consisting of the three parts mentioned, is called **semiotic**.

Semantics is also conceived, mostly by linguists, but also by some philosophers, as the theory of meaning. Although it is a fairly natural account, it creates a confusion pointed out by Quine (see Quine 1953a, p. 130):

> When we cleavage between meaning and reference is properly heeded […], the problems of what is loosely called semantics become separated into two provinces so fundamentally distinct as not to deserve a joint appellation at all. They may be called the *theory of meaning*

J. Woleński, *Semantics and Truth*, Logic, Epistemology, and the Unity of Science 45, https://doi.org/10.1007/978-3-030-24536-8_6

> and *the theory of reference*. 'Semantics' would be a good name for the theory of meaning, were it not for the fact that some of the best works in so-called semantics, notably, Tarski's, belong to the theory of reference. The main concepts in the theory of meaning, apart from meaning itself, are *synonymy* (or sameness of meaning), *significance* (or possession of meaning), and *analyticity* (or truth in virtue of meaning). Another is *entailment*, or analyticity of the conditional. The main concepts in the theory of reference are *naming*, *truth*, *denotation* (or truth-of), and extension. Another is notion of *values* of variables.

What Tarski himself understood by semantics is indicated in the following passages ((a) Tarski 1933, p. 252, (b) Tarski 1936, p. 401, (c) Tarski 1954a, p. 714):

> (a) […] we attempted to go further and to construct […] definitions and concepts belonging to semantics of a language – i.e. such concepts as satisfaction, denoting, truth, definability, and so on. A characteristic feature of the semantical concepts is that they give expression to certain relations between the expressions of language and the objects about which these expressions speak, or that by means of such relations they characterize certain classes of expressions or other objects. We could also say (making use of the *suppositio materialis*) that concepts serve to set up the correlation between names of expressions and the expressions of themselves.
>
> (b) The word 'semantics' is used here in a narrower sense than usual. We shall understand by semantics the totality of considerations concerning those concepts which, roughly speaking, express certain connexions between the expressions of a language and the objects and states of affairs referred to by these expressions. As typical examples of semantical concepts we may mention the concepts of *denotation*, *satisfaction*, and *definition*, […]. The concept of *truth* also – and this is not commonly recognized – is to be included here, at least in its classical interpretation.
>
> (c) the study of the relations between models of formal systems and the syntactical properties of these systems (in other words, the semantics of formal systems).

Perhaps the following picture might be offered in order to clarify some essential features of semantics as covering syntax, semantics in Tarski's sense, and pragmatics (see also Hiż 2004 on the relation of Tarski's semantics to grammar). Language plays a distinguished role in every part of semantics. We first proceed from simpler to more complex situations. We have language in itself, that is, a collection of expressions. To endow a collection of expressions with syntax, one must describe it from the point of view of admissible forms and relations between them. For example, if we say that an expression A occurs in (or is a part of) the expression $A \wedge B$, we make a statement about the syntax of the language of propositional calculus. Let us agree to call such statements 'syntactic'. Similarly, if we say that 'London' is the first word in the sentence 'London is the capital of the UK', we make a syntactic statement about ordinary English. In general, syntactic statements concern relations inside languages and are independent of semantic issues. In particular, a sentence is syntactically correct or not, independently of its truth-value. If one says that 'London' refers to London or that the form 'x is a prime number' is satisfied by the number 2, one utters a semantic statement, about English or arithmetic, respectively. Finally, pragmatic statements take into account the attitudes of users of languages, expressed by such words as 'asserting', 'asking' or 'guessing'. For example, the sentences 'Russell asserted that Wittgenstein was a genius' or 'Heidegger asked whether *Das Nichts nichtet*' represent pragmatic statements about English. Although borderlines between syntactic, semantic and pragmatic statements can be (and actually are) problematic in some cases (for instance, semantic and pragmatic factors sometimes correct syntactic errors), distinction of the three kinds of statements about language is out of question. Although these constatatations make no explicit reference to the concept of meaning, but semantics is very often characterized as the theory of meaning.

The concept of meaning is a source of continuous troubles for philosophers. On the other hand, it is difficult to imagine philosophical work ignoring the issue of what linguistic expressions mean. How to accommodate the concept of meaning in semantics? If we assume that the tripartite division of semantics *sensu largo* is exhaustive, only three possibilities remain: (a) the concept of meaning is added to syntax; (b) the concept of meaning is added to semantics *sensu stricto*; (c) the concept of meaning is added to pragmatics. The first possibility was attempted by the logical empiricists, but without success. Solution (b) requires either the concept of meaning as a new semantic primitive or its reduction to the concept of reference. It seems that supplementing the semantic vocabulary by the concept of meaning introduces dualism into semantics in the narrow sense, which is neither elegant nor easy to explain. On the other hand, the reduction of the concept of meaning to referential relations is at odds with the well-known fact that various and different intensions (meanings) can correspond to the same reference (extension). I opt for placing the concept of meaning in the territory of pragmatics. Roughly speaking, the meaning of expressions is their property that decides how they are understood in acts of communication. Of course, this is no definition. All known theories of meaning (mentalistic, behaviouristic, referential, subjective, objective, etc.) sooner or later agree that an understanding of expressions constitutes the main test for identifying of meanings. This also applies to Tarski's view, because in order to know how formal systems relate to their models, one needs understand the former.

(**DG1**) One could contest the account presented in the last paragraph by pointing out that I ignore intensional semantics based on the concept of intension as primitive, and favour the extensional theory based on the concept of extension. It is true that I understand semantics extensionally rather than intensionally. There are several serious reasons for that, but the main is that extensional semantics suffices for developing **STT**, at least for extensional languages. Whether this semantics suffices for intensional languages (for example, modal or epistemic) appears to be an open and controversial question, but I shall not deal with such languages (except in parenthetical remarks). Finally, one should also note that intensional semantics, even if it is indispensable in some cases, should be avoided so far where it is possible, because its ontological costs—consisting in the introduction of problematic intensional entities—seem to be too high.►

One can say that the situation is this. We start with syntax, then, referential relations are added and, finally, we pass to pragmatic issues. However, this perspective is not correct. Carnap's succession from pragmatics via semantics to syntax was not accidental. In fact, language, wherever it is used in communication, in mathematics or ordinary life, always manifests itself as a whole with pragmatics, semantic and syntactic dimensions. If we abstract from the users of language, that is, from the pragmatic dimension, referential (semantic) and syntactic relations remain. The next analytic abstraction ignores semantic matters and focuses on syntax. The result of this succession in abstracting is that the concept of meaning is just a pragmatic one, but that subsequent abstractions neutralize but do not ignore it. Thus, a very important consequence of what I said about the relations between syntax, semantics and pragmatics is that meaning functions as a presupposed attribute of expressions, although nothing is decided about its source. Thus, expressions appear as having a sense in virtue of a communicative tradition or more or less justified conventions—but always as meaningful. If so, we can overcome the gap observed by Quine between the theory of meaning and the theory of reference.

6.2 Historical Remarks on Semantic Terminology

Various difficulties concerning of how semantics could or should be understood as well as characterized are deeply rooted in its history. The term is, as expected, derived from a Greek word, namely *sema* (sign). However, Greek had also two other expressions, namely *semainein* (denote, designate, refer to) and *semantikos* (denoted, designated, referent, but also having meaning); thus, the contrast, to use contemporary way of speaking, between intensional and extensional aspects of meaning goes back to the very etymology of the semantic vocabulary. However, there is no accessible evidence that the denomination 'semantics' (or its counterparts in other languages) was in use before the end of 19th century. It was Michel Bréal (see Bréal 1897), the professor of comparative grammar in Collège de France who introduced the term *semantique*. For Bréal, semantics was a part of general linguistics concerned with the so-called lexical meanings of words and investigations of how such meanings change through time. Although Bréal's primary task was descriptive and historical, he did not shun more theoretical work. In particular, he looked at the laws of meaning-changes. He formulated, among other things, the rules of repartition or of preserving some features of lexicon, for instance, archaisms. Bréal's investigations undoubtedly created the origin a definite paradigm of research in linguistic semantics as focused of meaning changes and regularities in various ordinary languages.

Quine attributes the introduction of the word 'semantic' (as a noun) to Peirce (see Quine 1990a, p. 68):

> As used by C. S. Peirce, "semantic" is the study of the modes of denotation of signs: whether a sign denotes its object through causal or symptomatic connection, or through imagery, or through arbitrary convention, and so on. This sense of semantic, namely a theory of *meaning*, is used also in empirical philology: *empirical semantics* is the study of historical changes of meaning of words.

Quine's description well fits Bréal's account, because both stress the diachronic aspects of language. Peirce was more preoccupied with the logical aspects of signs than it had place in Bréal's case, I will not enter into details of Peirce's views on logic and language. Later linguists ascribed to semantics (sometimes called 'sematology') more ambitious aims than Bréal, and considered it as a part of a linguistic theory devoted to the study of the functions of language from a theoretical point of view.

(**DG2**) In Germany, the word *Semasiologie* was frequently employed as a label for the science of meaning. Semantics sensu *largo* is often identified with semiotic. In fact, Morris (see above) did not use 'semantics' but just 'semiotic' (derived from the Greek word *semeion* as other counterpart for 'sign'), so did Carnap (see the quotation above). Galen understood semiotic as diagnostic in medicine (inference about a sickness on the basis of its symptoms). This special use was generalized in the 17th century into a notion of semiotic as a general theory of signs. Locke is regarded as the classic exponent of this conceptual shift. According to him, semiotic investigates the nature of signs as related to things and helping to acquire knowledge. This shift moved semiotic from medicine to epistemology, and exposes the roots of the well-known ambiguity associated with the word 'sign'—which refers either to natural signs (for example, smoke as a sign of fire) or to proper signs (for example, words). *Semiologia* or even *semiologia philosophica* (Alexander Baumgartner) are other terms that should be mentioned in this context; also the works of Lambert became important in the development of semiotic. However, the modern uses of 'semiotic' owe their character to Peirce and Ferdinand de Saussure. These last two names are symbols that the contemporary work in semiotic has two relevant sources: philosophy and linguistics, and both cooperated in the development of semiotic and the relevant nomenclature.►

As far as philosophy and logic in the 20th century are concerned, the word 'semantics' was used only occasionally until the 1930s. For example, Ogden and Richards (see Ogden, Richards 1923, p. 2) mention a science of Semantics as dealing with relation between words and facts; they make vague reference to the work of Dr. Postgate (probably John P. Postgate, a classical philologist). That Quine (see his reference to Peirce quoted above) used the word 'semantic' as a noun, but not as an adjective, seems to provide evidence that there was no established use with regard to semantic matters. When Ramsey (see Ramsey 1925) considered antinomies in logic, he divided them into logical and epistemological. The latter were later baptized as semantic, and the former—as belonging to set-theory. Thus, Ramsey did not have the adjective 'semantic' for labelling the Liar Paradox and similar antinomies. Tarski (see Tarski 1930–1931, Tarski 1932) used the words *semazjologia* (Polish) and *Semasiologie*, and considered truth as a semasiological notion. This is an interesting example of how linguistic terminology influenced philosophical one.

Poland seems to be the first country in which the noun 'semantics' and the adjective 'semantic' (originally *semantyka* and *semantyczny* in Polish, but I will employ English terms in my reporting works of Polish logicians and philosophers) gained a wider philosophical popularity. The word 'semantics' occurs in Leśniewski 1927, p. 181. Kotarbiński (see Kotarbiński 1929, p. 20) spoke about semantics as concerned with the meaning-aspect of language. In the beginning of the 1920s, Leśniewski introduced in his lectures on the foundations of logic, the term 'semantic categories' for what Husserl understood by meaning-categories (*Beduetungskategorien*) (see Sect. 6.3). Ajdukiewicz discussed semantic functions, of which meaning is a paradigmatic example (see Ajdukiewicz 1931, p. 2). In 1930/31, the same author gave a course in Lvov devoted to selected problems of logical semantics (see Ajdukiewicz 1993). It seems that it was the first use of the term 'logical semantics' in the history (this is only a tentative hypothesis). In fact, Ajdukiewicz's discussed in his course mainly semantic categories (in Leśniewski's sense) and semantic antinomies, that is, syntactic rather (hence, the label 'syntactic categories' was used in Poland and elsewhere) than semantic (sensu *stricto*) problems (supplemented by remarks about the use of expressions), but the appearance of the term 'logical semantics' is interesting in itself. It was Tarski who introduced the word 'semantics' (see quotations in the previous section) in the meaning that was first accepted, and then became standard.

One other one tradition should be reviewed. I mean what happened in Vienna and logical empiricism. In the early 1930s, Carnap used *Semantik* as a synonym for *Metalogik* (see Carnap 1934, Carnap 1934a). He alludes to semantic matters only on the occasion of referring to Leon Chwistek's (a Polish logician) views. Carnap remarks (p. 9) that Chwistek's semantics is actually syntax just as *Metalogik*. Carnap was of course fully aware that of the linguistic notion of semantics as well as some other terminological proposals mentioned above like 'semasiology' or 'sematology'. He also used the hybrid term 'quasi-syntactic' for concepts that express relations between words and objects, but having complete syntactic translations. The word 'quasi-syntactic' and its meaning well displays Carnap's attitude to the effect that semantics (in Tarski's sense) is reducible to syntax. Carnap changed his mind under Tarski's influence (see the Introduction the present book, Coffa 1987). Carnap 1939 outlines (see above) the field of semiotic with its subdivision into three parts, and Carnap 1942 can be considered as the final stage of Carnap's journey from syntax to semantics (see Tuboly 2017 and Sect. 6.5); this book was the first comprehensive monograph on semantics *sensu stricto* in the entire history of logic.

6.3 Antinomies, Logical Types and Syntactic Categories

It is convenient to begin with Kurt Grelling's antinomy. Adjectives can be divided into autological and heterological. An adjective is autological only if it has the property expressed by it. For example, the adjective 'short' is short because it has the property expressed by the predicate 'is a short word'. On the other hand, the adjective 'heterological' is not short by any typically accepted standards of assessing the length of words. Now consider the following question:

(1) Is ‘hererological’ autotological or heterological?

First, assume that ‘heterological’ is autological. By definition, this adjective possesses just the property that is expressed by it. Consequently, the word ‘heterological’ must be counted as heterological, and we have

(2) If ‘hererological’ is autological, then it is heterological.

Now assume that ‘heterological’ is heterological. Consequently, ‘heterological’ does not have the property of being heterological, expressed by it and must be autological. That gives

(3) If ‘heterological’ is heterological, then it is autological.

Putting (2) and (3) together, leads to

(4) ‘Heretological’ is autological if and only if ‘heterological’ is heterological,

which appears as internally inconsistent. This antinomy belongs to semantics, because it involves reference to denoted properties.

Apart from semantic antinomies, we also have set-theoretical paradoxes, which arose in naïve (or Cantorian) set theory—of which the puzzle, discovered by Cantor himself—of the set of all sets is the simplest one. Intuitively, the term ‘the set of all sets’ refers to the largest set, because it contains all possible sets. Denote this set by the symbol $\oplus$. By a theorem of set theory about the relation between any set $\mathbf{X}$ and the set $\mathbf{2}^{\mathbf{X}}$ of all its subsets, we have

(5) $\oplus \subset \mathbf{2}^{\oplus}$.

This means that the set $\oplus$ is smaller than the set of all its subsets. This paradox can be resolved by the observation than the set $\oplus$ does not exist, since a theorem of set theory precludes its existence. The situation is similar to some extent to the fate of the traditionally accepted principle that the part must be smaller than the whole of which is a part. When Galileo showed that subsets of the set of natural numbers can be equal to the whole set of natural numbers, it showed that the intuition behind the accepted part/whole principle was erroneous. The reasoning about $\oplus$ and $\mathbf{2}^{\oplus}$ shows that the same concerns the set of all sets, because its mathematical treatment goes against preliminary intuitions.

The famous antinomy, discovered by Russell in 1902 cannot be resolved in this simple way. The Russell Paradox runs as follows. Divide all sets into normal and abnormal (or non-normal; this division is exclusive and exhaustive). The former are not their own elements, the latter are. Most sets are normal. For example, the set of cities is not a city, the set of animals is not an animal, etc.—hence, the adjective ‘normal’ in this context. On the other hand, every set of sets is a set and thereby is abnormal. Consider the set $\otimes$ defined by the condition:

(6) For any $x \in \otimes$, x is normal.

That means that is the set of all sets which are not elements of themselves. Now let us check whether $\otimes$ is itself is normal or abnormal. Assume that $\otimes$ is normal. This implies that $\otimes \notin \otimes$. But by definition, we have that $\otimes \in \otimes$. Thus, it is abnormal, if $\otimes$ is normal. If we now assume that $\otimes$ is abnormal, then $\otimes \in \otimes$, but, once more by definition, $\otimes$, as an element is itself, is normal as element of itself. Hence, $\otimes$ is normal, if it is normal. Both assumptions give that $\otimes$ is normal if and only if $\otimes$ is abnormal, symbolically.

(7) $\otimes \in \otimes$ if and only if $\otimes \notin \otimes$,

but this formula is paradoxical, because inconsistent.

The idea of the set of all normal sets cannot be rejected in the same manner as employed in the case of the set of all sets, because no principle or theorem of Cantorian set theory precludes the existence of the set $\otimes$. Russell solved the problem by his famous (simple) theory of logical types. This theory divides all objects into definite logical types. Omitting relations and concentrating exclusively on sets, we have that individual objects form the type 1, sets of individuals belong to the type 2, sets of sets populate the type 3, and so on. If a set is of type n, its elements have type $n - 1$. Now the theory excludes sets that whereby the type of the elements of the set equals the type of this set; we also say that the theory of logical types precludes the existence of such sets. For example, objects designated as $\otimes$ in (5) and (6), cannot be their own elements, for they belong to different logical types. The paradox disappears because the set of all normal sets does not exist.

(**DG3**) The resolution of paradoxes can be achieved in axiomatic set-theory. It means that "dangerous" sets are excluded by special axioms. I consider the Zermelo–Fraenkel set theory (see Fraenkel, Bar-Hillel, Lévy 1973 for an extensive treatment of various systems of set theory and ways of solving antinomies) and restrict my remarks only to this system. The naïve scheme of comprehension (the principle of naïve set theory) says that any condition defines a set. In order to exclude paradoxical sets, a refined axiom (of comprehension) is adopted, namely, that if **X** is a set, any subset of **X**, is also a set (accordingly, the concept of set is primitive and must be characterized by axioms; I skip how it is done). Since the collection of all sets has no overset, the new comprehension scheme does not apply to it. Consequently, the object $\otimes$ is not a set. One can accept the existence of $\otimes$ as a specific object (such objects are sometimes identified as too big to be sets). Although the original theory of logical types is nowadays considered as obsolete, its main idea—namely, that of stratification of all objects into levels is preserved, to some extent, by the so-called cumulative hierarchy of sets. It assumes that sets are constructed by stages. At stage 1 we have individuals, then sets of individuals at stage 2, and so on. We can eventually dispense with individuals and work only with sets, according to the rule that everything is a set. Let me add that the main arguments against Russell's theory are directed against its ramified version, which stratifies objects without particular types into orders, but I will not enter into details of this approach—except to point out that these complications were introduced mainly with the aim of solving some mathematical problems. Incidentally, one can observe that the set of all abnormal sets does not pose any problems.►

(**DG4**) Two meanings of the word 'paradox' have to be sharply distinguished. Firstly, this word refers to an inconsistency (contradictions), like in the case of the difficulties discovered by Cantor, Russell, and Grelling (as well as many other authors). Such inconsistencies are also called logical antinomies. Recall that this group of antinomies was divided (see Sect. 6.2) into logical and epistemological (semantic), but it does not change the situation that both kinds are provable as contradictions. Secondly, paradoxes are also unexpected or strange (according to some assumed criteria) conclusions, for instance, the Twins paradox in special relativity or the Banach–Tarski paradox with regard to the decomposition of a sphere (this result is obtained via the axiom of choice). Also Frege's puzzle (see Chap. 4, Sect. 4.2) belongs to this group. Note that I use the term 'paradox' as a synonym for 'antinomy'.►

Russell's account of logical types can be interpreted in two ways—namely, either ontologically or linguistically. While the former view considers types as classes of objects, but the latter sees them as classes of expressions. According to the second interpretation, the expressions '$\oplus \in \oplus$' and '$\otimes \notin \otimes$' are ill-formed. Here we have the point at which the theory of logical types meets the theory of syntactic categories (recall that they are also called 'semantic categories'). Consider the sign referring to the being an element of a set as the two-placed function forming a sentence (or a sentential formula with a free variable) from two arguments. More specifically, we have the predicate '… is an element of a set ——'. According to rules of the theory of logical types in its linguistic interpretation, expressions taken to fill the places indicated by '…' and '——' have to be of different categories. If the symbol that replaces the place '…' is of category n, the symbol replacing the place '——' must be of the category $n + 1$. Generally speaking, the names of sets are always of a higher type than the names of their elements. For example, in the sentence.

(8) London belongs to the set of cities,

the expression 'London' belongs to the category of proper names, but the term 'the set of cities' has another status (I skip a more detailed qualification, because I do not like to enter the problem of definite descriptions). A much clearer picture emerges when we rewrite the statement (8) as

(9) London is a city.

We can look at (9) as a first-order sentence, in which 'London' is a proper name and 'is a city'—as functioning as a one-place predicate expression. This reformulation leads to a distinction between two syntactic categories of nominal (in the intuitive sense) expressions—namely, proper names and predicates. However, the expression '$\mathbf{X} \in \mathbf{X}$' is meaningless under both readings.

(**DG5**) So far I said nothing about how semantic antinomies can be resolved. Clearly, the theory of types does not suffice in this respect, because we have no obvious reason to consider 'autological' and 'heterological' as words belonging to different logical types. Leśniewski and Tarski proposed that the proper solution of semantic antinomies should be based on the object-language/metalanguage distinction (in what follows I use the term ‘language/metalanguage distinction’). This distinction is absolutely fundamental for **STT**, but I postpone details until Chap. 7, Sect. 7.4.►

The important problem concerns whether there exists the logical type comprising all other types. A positive answer to this question is tempting, particularly for most (I guess) philosophers, who might suggest a positive answer pointing out that being *qua* being has highest type, but we must say “no”, because the “universal” type would just as inconsistent as the set of all sets. Another serious problem that arises in the theory of logical types is how to interpret such concepts as that of logical constant, set, relation, etc. They appear at every level of the hierarchy of types. For example, is meaning of ‘and’ the same in the framework of different type or not? Are meanings of Russell was fully aware of this issue and proposed the idea of typical ambiguity, that is, the solution that although the meaning of related concepts is practically the same at every level, although they are formally different because they apply to objects populating different types. Strictly speaking, we should always say, for example, the set of all objects of the type n; thus, the phrase ‘the set of all objects’ is incomplete without indicating that we are speaking about a concrete type. Tarski (see Tarski 1986a), following to some extent the Erlangen program in geometry to some extent, suggested that some concepts, deserving to be called logical, should be interpreted as invariants with respect to one-to-one transformations. He argued that the concepts of set and its cardinality, relation, identity, and difference belong, to this group. This actually solves the problem of typical ambiguity for the mentioned notions and contributes essentially to logicism in the foundations of mathematics, because it enables us to regard the concept of set as a logical notion. Yet this problem is not of particular importance for first-order axiomatic set theory. The reason for this is that the latter is an extension of elementary logic, and this extension can hardly be regarded as a part of **FOL** (see arguments given in Chap. 5, Sect. 5.4).

Another strong objection against the theory of logical types is raised by the question of whether it can be formulated without violating its own principles. Clearly, the theory in question says something about all logical types. However, if so, it must contain expressions referring to all types, but that is impossible in the light of the principles the theory in question. The difficulties perhaps can be summarized as follows. The theory of logical types imposes some requirements on concepts and their admissible combinations. It asserts thereby something about all concepts, in particular, that some of their combinations are incorrect. Now, if the theory of types asserts something about all concepts, it applies to its own conceptual apparatus. On the other hand, according to this theory, concepts cannot be self-referential, that is, referring to themselves. Combining all features of the theory of logical types, we find that it cannot be true, because it is meaningless, according to its own standards. To put it in another way, language that asserts something about all logical types (interpreted linguistically) is impossible, because it would violate the principles of the simple theory of types (*a fortiori*, the ramified theory too). Since analysis of the theory of logical types is not my main target in this book, I will not discuss proposed ways to solve the difficulty mentioned in this paragraph (see Copi 1971 for a survey). However, two remarks are in order. Firstly, the similarity of the problem of the type of all types to the problem of the set of all sets is evident. Secondly, as we will see in Chap. 7, Sect. 7.4, the extension of these arguments to the hierarchy of languages determined by the language/metalanguage distinction is straightforward.

Roughly speaking, two or more expressions belong to the same syntactic category if and only if they can be mutually substituted in an expression *E* without violating the syntactic correctness of this expression. Ajdukiewicz (see Ajdukiewicz 1935) proposed a very simple algorithm to check the syntactic correctness of complex expressions. Without treating it in its full generality, let me explain it by means of concrete examples. We divide expressions into sentences, proper names and functors. Every category has its own symbolic index. Sentences -are indexed by the letter ***s***, and names by the letter ***n***. The indexes of functors (for example, connectives of **PC**) are fractions where the numerator indicates the kind of expressions formed by a given functor and the denominator refers to the kind and number of arguments. For example, the functor ‘and’ acting as the conjunction (n the sense of **PC**) has the index ***s/ss***, because it forms a sentence from two other sentences. We interpret *n*–place predicates as functors which form sentences from *n* proper names. For example, the predicate ‘is older than’ forms a sentence (or an open formula of the sentential category) from two proper names. Having a concrete sentence, we can parse (decompose) it into its syntactic components by writing their indices (the succession is determined by the structure of the given expression). Thus, the formula ‘*p* and *q*’ is parsed into **<*s*, *s/ss*, *s*>**, but the sentence ‘*a* is older than *b*’ into **<*n*, *s/nn*, *n*>**. The rule of correctness states that a sentence is correct (well-formed) if and only if quasi-arithmetical simplifications culminate it ***s*** as the final result. It is easily to see that both of the mentioned decompositions yield ***s*** as the final outcome. On the other hand, the expression ‘*a* is older than’ is not a correct sentence, because it is parsed into **<*n*, *s/nn*>**–and finally, ***s/n***.

Consider the sentence (9) once again. Its parsing via syntactic categories gives the sequence **<*n*, *s/nn*, *n*>** and shows that (9) is correct. This analysis follows Polish grammar in which ‘London’ and ‘a city’. The Polish counterpart of (9) is

(10) *Londyn jest miastem*

where *Londyn* (London) and *miastem* (a city; Polish has endings indicating grammatical cases of nouns, but this circumstance is not relevant here) are considered as names—*Londyn* as proper (individual) and *miastem* (a general one); this analysis follows Polish grammar; Polish language has no articles. Another parsing is determined by treating ‘is a city’ as a predicate. Its index is ***s/n***. Under this syntax, the sequence related to (10) has the form **<*n, s/n*>** and is correct as well. This example shows that Ajdukiewicz’s algorithm is (relatively) independent of the syntactic peculiarities of particular languages. However, if we take into account the grammar of **FOL**, references of individual constants (proper names) and references of predicates are of different logical types, because the former are individuals, but the second—sets. On the other hand, the parsing that are based on Polish grammar treats the denotations of ‘London’ (*Londyn*) and ‘a city’ (*miasto*) as belonging to the same logical type. Yet ‘is’ has different indices in both cases, namely ***s/n*** and ***s/nn***, respectively. An interesting fact is that assertions about entities belonging to different types can be syntactically correct. We encounter here a considerable advantage of the theory of syntactic categories over the theory of logical types.

6.4 General Historical Remarks on the Development of Semantics in the 20th Century

Although the term ‘semantics’ did not appear in philosophy until fairly recently (see Sect. 6.2), dealing with several semantic problems has a much longer tradition (see Coffa 1991 for the post-Kantian period), because language was always felt as something important philosophical enterprise. As far as logical semantics is concerned many relevant ideas were anticipated by Bolzano—in particular, his account of the concept of semantic consequence and truth or the treatment of analytic sentences. Frege’s distinction of *Sinn* and *Bedeutung*, Russell’s analysis of definite descriptions, his (as well as other authors) work on various paradoxes (Russell’s theory of types has a clear semantic dimension) or Wittgenstein’s picture theory of meaning, appear as splendid examples of considerations and achievements in logical semantics. The initial work in model theory (in logic) due to Ernst Schröder, Löwenheim and Skolem (can be also included here. Russell had an idea of the completeness property of a logical system when he said in *Principia* (see Whitehead–Russell 1910–1913, p. 12) that an adequate logic should allow us to deduce all logical truths. He also stressed the general philosophical significance of studying language (Russell 1903, p. 42):

> The study of grammar, in my opinion, is capable of throwing far more light on philosophical questions than is commonly supposed by philosophers.

This proclamation is crucial for understanding the heart of analytic philosophy, although the term ‘grammar’ can be misleading—Russell had in his mind logical grammar, not the lied belonging to linguistics.

However, these particular ideas, treatments or results did not lead to a general semantic theory or even an idea of how to do it. Jaakko Hintikka (see the essays in Hintikka 1997) and Martin Kusch (see Kusch 1989) tried to explain this historical fact by a distinction between language as calculus (**LCA**) and language as universal medium (**LUM**), which is related to van Heijenoort’s distinction between logic as language and logic as calculus (see van Heijenoort 1985 and the preceding chapter in this present book). We have two sets of ideas (see Kusch 1989, pp. 6–7):

LUM	LCA
(I) Semantics is not accessible	Semantics is accessible
(II) Different systems of semantic relations are inconceivable	Different systems of semantic relations are conceivable
(III) Model theory is rejected	Model theory is accepted
(IV) Semantic Kantianism is adopted	Semantic Kantianism is rejected
(V) Metalanguage is illegitimate	Metalanguage is legitimate
(VI) Truth as correspondence is not intelligible	Truth as correspondence is intelligible
(VII) Formalism is linked with the thesis that semantics is not accessible	Formalism is linked with the thesis that semantics is accessible

For Kush and Hintikka, who defend the **LCA** account, language should be considered is a re-interpretable set of expressions. Consequently, language, on this view, is not conceived as a purely formal construction, and thereby, non-interpreted (or pure) calculus. On the other hand, according to **LUM** theory of language, semantics understood generally as comprising the whole of semantic relations in a language, is simply ineffable, that is, cannot be expressed by any comprehensible linguistic devices. As we see from the above comparison of the two accounts, they are precisely contrary on every point in the list (I)–(VII). The fate of the correspondence theory of truth may provide a stark contrast of the fundamental difference between **LUM** and **LCA**.

Hintikka and Kusch use the **LUM/LCA** distinction as a device for displaying a map of the contemporary philosophy of language. Frege, Russell and Wittgenstein represented the **LUM** account, but Husserl, Gödel and Tarski accepted the **LCA** view. In particular, the former perspective blocked the development of semantics, because it forces us to look at language as the reality that precludes any way of speaking about itself. This way of looking at the history of contemporary semantics is certainly illuminating, but it also provokes some objections. In particular, it seems think to me that in the spirit of **LCA** we are able to distinguish both a global and a local approach. The latter can be consistent with the **LUM** account to some extent. For example, one can consider fragments of language as calculi, and the whole of it as the universal medium. It seems that in some cases Frege was a localist in some cases and a globalist universalist as far as the full scope of language is concerned. I also have doubts whether Gödel accepted the **LCA** approach in all details, because he looked for a universal system of set theory—at least in the last years of his scientific carrier. Thus, except for some hermeneuticians, like Heidegger or Gadamer, only Wittgenstein (at least, in the *Tractatus*) remains as a pure universalist, although his idea that propositions represent facts contributed to semantics—perhaps against his intentions, since he rejected the view that one can formulate menaingful propositions on relations holding between what is represented and what representation is. It seems that the characterization of **LCA** should be supplemented by an observation that language is, according to this perspective, stratified into infinite hierarchy of levels, re-interpretable at every stage. Tarski embraced this move.

Another point of disagreement concerns Husserl. It is true that some of his views influenced the development of formal semantics in the 20th century. For example, Bar-Hillel (see Bar-Hillel 1970a) maintains that Husserl's distinction between *Unsinn* and *Widersinn* anticipated Carnap's pair formation/transformation rules as generating what is meaningful and what is not. According to Bar-Hillel's interpretation, Husserl's idea of pure grammar might have been a pattern for Carnap's formal syntax, but this conjecture has no textual basis, at least in sources accessible for me. Surely, Husserl (see Husserl 1984, p. 33) anticipated many particular points important for semantics, for instance, he had a fairly clear idea of semantic consequence. Further, Husserl's idea of a manifold (see Husserl 1929, 28–32) as a domain of elements defined by a set of axioms is similar to the concept of a class of models generated by an axiomatic system.

In his contribution, Kusch remarks (Kusch 1989, p. 60):

> Concerning investigations three and four [in Husserl's *Logical Investigations* – J. W.] the first noteworthy fact relates to the historical role these studies played in the development of semantical approaches in formal logic. As is well-known, the main gate through which these ideas entered modern logic was the work of Tarski and other Polish logicians. […] it is remarkable that it was Husserl's *Logical Investigations* and especially the third and fourth investigation that exerted a strong influence in Warsaw between the two world wars. This influence has been described as comparable to the influence of Wittgernstein's *Tractatus* in the Vienna of the twenties and thirties. […]. It can be considered as indirect evidence for attributing to Husserl the calculus conception of that a precise formal semantical theory was developed where his influence was the strongest. And this influence did not only remain an abstract, unspecified level. […]. Ajdukiewicz's and Leśniewski's seminal work on categorial grammar had as its starting point Husserl's fourth investigation concerning the ideal logical grammar.

The idea of semantic categories was no doubt very influential in Poland. It is clearly documented by references in the works of Ajdukiewicz, Leśniewski and Tarski. For example, Leśniewski wrote (Leśniewski 1929, pp. 421–422):

> In 1922 I outlined a concept of semantical categories as a replacement for the hierarchy of types, which is quite unintuitive to me. Frankly, I would still today feel obliged to accept this concept even if there were no antinomies at all. From a formal point of view my concept of semantical categories is closely related to the well-known type theories […] especially with regard to their theoretical consequences. Intuitively, however, the concept is more easily related to the thread of tradition running through Aristotle's categories, the parts of speech of traditional grammar, and Husserl's meaning categories.

Although Leśniewski explicitly mentioned Husserl, I do not agree that influence of the latter essentially contributed to the development of semantics in Poland as far as the general point is concerned, in particular, related to the **LCA**-idea of language. A more important factor consisted in the acceptance by Polish philosophers and logicians of Brentatno's view that mental phenomena are intentional. Since according to most Polish philosophers, mental phenomena manifest themselves through language, the latter is also intentional in the sense that expressions (at least those that serve as sentences and names) have links to the world. If Husserl was important in this respect, it was owing to the fact resulted from the fact that he too represented this way of thinking about language and the world, although one could also observe that Husserl's later philosophy turned toward the **LUM** account due to its transcendental character. But before he passed to transcendental phenomenology he certainly contributed to a general philosophical climate, which turned out to be favourable for semantics. Let me illustrate this point by quoting a fragment from Leśniewski (Leśniewski 1929, p. 487/478):

> Having no predilection for 'various mathematical games' that consist in writing out according to one or another conventional rule various more or less picturesque formulae which need not be meaningful, or even – as some of the 'mathematical gamers' might prefer – which should necessarily be meaningless, I would not have taken the trouble to systematize and […] check […] the directives of my system, had I not imputed to its theses a certain specific and completely determined sense, in virtue of which its axioms, definitions, and final directives […] have for me an irresistible intuitive validity. I see no contradiction, therefore, in saying that I advocate a rather radical 'formalism' in the construction of my system even though I am an obdurate 'intuitionist'. Having endeavoured to express some of my thoughts on various particular topics by representing them as a series of propositions meaningful in various deductive theories, and to derive one proposition from others in a way that would harmonize with the way I finally considered intuitively binding, I know no method more effective for acquainting the reader with my logical intuitions than the method of formalizing any deductive theory to be set forth. […] theories under the influence of such formalization cease to consist of genuinely meaningful propositions which […] are intuitively valid. But I always view the method of carrying out mathematical deduction on an 'intuitionistic' basis of various logical secrets as a considerably less expedient method.

Independently of whether the **LUM/LCA** distinction provides an adequate framework for a philosophical explanation of the development of semantics, several additional concrete circumstances should be also mentioned as fairly significant for this process. We must also remember that the acceptance of semantics had to rebut various challenges. Tarski diagnosed the situation in the middle of the 1930s in the following passage (Tarski 1936, p. 401):

> Concepts from the domain of semantics have traditionally played a prominent part in the discussions of philosophers, logicians and philologists. Nevertheless they have long been regarded with […] scepticism. From the historical point of view this scepticism is well founded; for, although the content of the semantical concepts, as they occur in colloquial language, is clear enough, yet all attempts to characterize this content more precisely have failed, and various discussions in which these concepts appeared and which were based on quite plausible and […] evident premises, have often led to paradoxes and antinomies. It suffices to mention here the antinomy of the liar, the Grelling–Nelson antinomy of heterological terms and the Richard antinomy of definability.

Thus, for Tarski, the danger of antinomies caused the scepticism toward semantics. According to the traditional view, all paradigms in logic and the foundations of mathematics, logicism, formalism and intuitionism, arose as definite replies to challenges stemming from antinomies and the difficulties caused by them. Although there is not denying that antinomies—in particular the Russell Paradox of the set of all sets which do not belong to themselves (see below)—provided a serious impetus for intensive logical research—one could argue that inconsistencies did not make a big impression on Brouwer, Hilbert or Zermelo. It is not easy to give an answer how antinomies were relevant for the development of logic and the foundations of mathematics. They undoubtedly inspired Russell's logicism. The theory of logical types (see below) came about and was improved in order to provide a weapon against paradoxes. However, one may argue that Brouwer (intuitionism) and Hilbert (formalism) had other reasons for developing their foundational projects; as far as antinomies are concerned, Brouwer pointed out that difficulties in the foundations of set theory resulted from the unrestricted use of the principle of excluded middle, but, on the other hand, Zermelo's axiomatization of set theory excluded "dangerous" sets.

To complete this survey of the development of semantics in the 20th century, I will briefly report the views of Gödel and Tarski (see Chap. 8, Sect. 8.5 for a more extensive analysis and comparisons). In his earlier works, Gödel, (see Gödel 1930, Gödel 1931) considered semantic concepts, particularly the concept of truth, as merely auxiliary and heuristic. More specifically, he wanted to eliminate semantic categories in favour of syntactic notions. For Gödel, particularly in his later remarks (some of them were not published during his life time; see details in Chap. 8, Sect. 8.5) the general philosophical context of the 1920s and 1930s largely prevented a serious treatment of the concept of truth. More specifically, Gödel associated this trend with the negative attitude toward non-finitary methods of reasoning, represented by leading logicians like Skolem, Hilbert or Jacques Herbrand. The same can be said about the consequences of logical empiricism and its syntacticism toward language (see the next section). One could say that Gödel himself was a victim of the same attitude, although he, at least as a philosopher, did not share it. Most commentators diagnose that his "caution" (as Solomon Feferman calls it; see Feferman 1984) concerning the concept of truth was partly influenced by his scientific environment. Gödel worked on Hilbert's program and his proof of the completeness theorem was a very important step on the way to finitary metamathematics; although the incompleteness theorems went in the exactly opposite direction, and appeared to be undesirable by-products, they belonged of the same (formalistic) foundational project. Gödel never accepted most views of the Vienna Circle, but he grew up in the climate of this philosophy, and first works in logic clearly show that formalism functioned as his main guide. Even if his private opinions in the philosophy of mathematics were more Platonic than he publicly admitted, he fully agreed that proofs in metamathematics should be constructive and arithmetized. At any rate, various objections to semantics in the first quarter of the 20th century resulted from thinking focused on the concept of proof rather than on the notion of truth and considering the former as the basic category of metalogic.

Tarski functioned in a philosophical as well as mathematical atmosphere that favoured semantics. Łukasiewicz, Leśniewski and Kotarbiński, Tarski's main teachers in logic and philosophy, studied with Twardowski and inherited basic Brentano's views about mind and language (see the preceding section). Yet it is not easy to assess the extent to which Tarski was influenced by this tradition, because he did not like to explain his own philosophical views (Mostowski 1967, p. 81; see also Mycielski 2004):

> Tarski, in oral discussions, has often indicated his sympathies with nominalism. While he never accepted the 'reism' of Tadeusz Kotarbiński, he was certainly attracted to it in the early phase of his work. However, the set theoretical-methods that form the basis of his logical and mathematical studies compel him constantly to use the abstract and general notions that a nominalist seeks to avoid. In the absence of more extensive publications by Tarski on philosophical subjects, this conflict appears to have remained unresolved.

Keeping in mind this opinion of a person especially competent to speak about Tarski and his views, let me speculate a bit (see also Woleński 1995a). It is improbable that Tarski was completely neutral with respect to some general inclinations of his teachers (I neglect here various concrete influences—for example, nominalism, empiricism, and reism). All of his scientific activity documents that he considered, as did Łukasiewicz, every formal problem as worthy of being investigated without any philosophical prejudices. On the other hand, like Leśniewski, Tarski had a very strong feeling for philosophical aspects of formal work and to Leśniewski's view that logic successfully codifies intuition (see Introduction). Thus, if the concept of truth had interesting intuitive, formal, and philosophical aspects there was no reason to exclude it from strict logical research.

Perhaps Polish mathematical legacy acquired still greater importance for Tarski's prosemantic attitude. The Polish mathematical school with its concentration on set theory, topology, and their applications in other branches of mathematics had no methodological prejudices concerning admissible methods (Sierpiński 1965, p. 94; Sierpiński held this view since 1918 at least):

> Still, apart from our personal inclination to accept the axiom of choice, we must take into consideration […] its role in the Set Theory and in the Calculus. On the other hand, since the axiom of choice has been questioned by some mathematicians, it is important to know which theorems are proved with its aid and to realize the exact point at which the proof has been based on the axiom of choice; for it has frequently happened that various authors have made use of the axiom of choice in their proofs without being aware of it. And after all, even if no one questioned the axiom of choice, it would not be without interest to investigate which proofs are based on it and which theorems can be proved without its aid – this, as we know, is also done with regard to other axioms.

Tarski himself fully concurred with this standpoint (Tarski 1962, p. 124):

> We would of course fully dispose of all the problems involved [that is, concerning inaccessible cardinals – J. W.] if we decided to enrich the axiom system of set theory by including (so to speak, on a permanent basis) a statement which precludes the existence of 'very large' cardinals, e.g., by a statement to the effect that every cardinal $> \omega$ is strongly incompact. Such a decision, however, would be contrary to what is regarded by many as one of the main aims of research in the foundations of set theory, namely, the axiomatization of increasingly large segments of 'Cantor's absolute'. Those who share this attitude are always ready to accept new 'construction principles', new axioms securing the existence of new classes of 'large' cardinals (provided that they appear to be consistent with old axioms), but are not prepared to accept any axioms precluding the existence of such cardinals – unless this is done on a strictly temporary basis, for the restricted purpose of facilitating the metamathematical discussion of some axiomatic systems of set theory.

As far as metamathematics is concerned, Tarski summarized the situation in the following statement (Tarski 1954, p. 713):

> As an essential contribution of the Polish school to the development of metamathematics one can regard the fact that […] it admitted into metamathematical research all fruitful methods, whether finitary or not.

And here we have the point where semantics and metamathematics meet, so to speak, on the official level. It is, of course, not true that truth did not belong to mathematical jargon before Tarski. It was used informally, although essentially—when the completeness problem was being formulated, for example. However, as Gödel remarked, it was not formally elaborated, although he employed it himself in stating the completeness theorem (see Gödel 1930), and in informal explanations the first incompleteness theorem (see Chap. 8, Sect. 8.4). It was Tarski, among others, who quite consciously combined philosophical and mathematical interest in the precise elaboration of semantic concepts. It was accidental by no means that the first steps of the formal analysis of truth (the concepts of satisfaction and definability) appeared in a paper (see Tarski 1931) in which some problems of descriptive topology were studied—a domain in which the transfinite plays an important role.

(**DG6**) Robert Vaught (see Vaught 1974, p. 161) reports Tarski's dissatisfaction—expressed in his seminars in 1926–1928—that the concept of satisfaction was not properly defined. This explains a point raised by Georg Kreisel (see Kreisel 1987, p. 122):

> According to Andrzej Mostowski, in a conversation in Tarski's presence, the latter and his students had no confidence in Gödel's paper [Gödel 1930 – J. W.] when they saw the relevant issue of the *Mhfte* [an abbreviation for the journal in which the paper appeared – J. W.] in Warsaw. Why? Gödel had not formally defined validity! Anybody who is surprised by this knows *ipso facto* that he simply had no feeling for the subject."

The ironic tenor of Kreisel's remark is not justified. That Gödel, who was a great authority and perhaps the most influential logician after Aristotle, "had no feeling for the subject" (by the way, this opinion does not seem quite correct; see Chap. 8, Sect. 8.7) is no sufficient reason to consider Tarski's way of thinking as odd and blame it. Both attitudes, Gödel and Tarski's, had their own definite roots and outcomes. In particular, Tarski, educated in a very demanding environment as far as the clarity of concepts having central importance for given investigations is involved, must have had nurtured need to look for satisfactory definitions, in particular, of satisfaction and truth.►

6.5 The Triumph of Semantics: Carnap's Case

I mentioned how the term 'semantics' heavily functioned in the works by logical empiricism. Since this movement played an essential role in the adoption of semantic methods in philosophy, the evolution of the attitude of this movement toward semantics is particularly instructive. Carnap's seminal book on the logical syntax of language (Carnap 1934) still represented syntacticism, that is, the view that all genuine metalogical problems are syntactic in character and should be investigated by methods belonging to syntax. He changed his attitude under Tarski's influence (see Introduction to this book). In fact, Carnap had two important conversions—the first, to metalogic, and the second to semantics. Carnap's turn to metalogic was a revolt against Wittgenstein, who, according to his **LUM** idea of language, considered all statements about the logical form of language or its relation to the world as plainly nonsensical. In contradistinction to this opinion, Carnap, under the influence of Hilbert and Gödel, admitted assertions about syntactic properties of expressions as unconditionally meaningful. However, this view does not address whether such statements about syntax are syntactic at all.

In order to apply logical syntax in metaphilosophy, Carnap introduced three related distinctions:

(I) Objectual questions versus logical questions;
(II) Meaning question versus formal questions;
(III) Material mode of speech versus formal mode of speech.

Distinction (I) was introduced by Carnap in the following way (Carnap 1934, p. 277):

> The questions dealt with in any theoretical field – and similarly the corresponding sentences and assertions – can be roughly divided into *object-questions* and *logical questions*. (This differentiation has no claim to exactitude; it only serves as a preliminary to the following non-formal and inexact discussion.) By object-questions are to understood those that have to do with the objects of the domain under consideration, such as inquiries regarding their properties and relations. The logical questions, on the other hand, do not refer directly to the objects, but to sentences, terms, theories, and so on, which themselves refer to objects. (Logical questions may be concerned either with the meaning and content of the sentences, terms, etc., or only with the form of these […]).

For example, the sentences 'New York is a big city' and '10 is an even number' express answers to object-questions and are object-sentences. On the other hand, the sentences '*A* and *B* are equivalent', '*B* and *C* are synonymous' or '*B* is entailed by *A*' belong to the realm of logical questions.

Carnap's fundamental claim concerning (II) is that all questions of meaning are reducible to formal questions, that is, sentences about meaning are replaceable by syntactical sentences. For instance, the content of an extralogical (synthetic) sentence is the class of its non-analytic consequences. This immediately entails that tautologies (analytic sentences) are devoid of any content (they have no synthetic content). Furthermore, the statement that a proposition asserts something necessary means that it is a tautology, and the statement that two propositions are compatible means that they are consistent. Combining (I) and (II), we can distinguish two different kinds of sentences: object-sentences and syntactical sentences. However, this distinction is not sufficient to diagnose what philosophical sentences are. This is due to the fact that philosophy, particularly, in its traditional guise covers different kinds of problems. Sometimes it considers genuine empirical problems, for example, psychological. Logical questions are another subject of philosophical research. Of course, we can delegate the results of these investigations to science.

Yet we encounter in philosophy statements, which cannot be classified as belonging either to pure logic or empirical science. Consider the sentence ‘Five is not a thing, but a number’. It is usually intended as an object-sentence. Its peculiarity consists in having the word ‘thing’ being employed by philosophers to indicate something concerning the ontological status of numbers. Carnap called such sentences pseudo-object-sentences and tried to explain their status. In order to do it, he introduced so called quasi-syntactic sentences. Here is an explanation (Carnap 1934, pp. 233–234):

> Let B be a domain of certain objects whose properties are described in the object-language S_1. Assume that there exists in reference to B an object-property E_1, and in reference to S_1 a syntactical property of expressions E_2, such that always and only when E_1 qualifies an object, E_2 qualifies the expression which designates that object. We shall call E_2 the syntactical property correlated to E_1. E_1 is then a property which is, so to speak, disguised as an object-property, but which, according to its meaning, is of a syntactical character (or sometimes a pseudo-object-property). […]. A sentence which ascribes the property E_1 to an object **c** is called a quasi-syntactical sentence; such a sentence is translatable into the (proper) syntactical sentence which ascribes the property E_2 to a designation of **c**.

Carnap successfully applied his syntactic machinery to several metalogical problems. He distinguished two languages, namely Language I (roughly speaking, first-order language) and Language II (roughly speaking, higher-order language). The latter was introduced in order to meet the challenge (stated by Gödel in oral discussions) that the concept of analyticity cannot be defined in Language I, because we have true undecidable arithmetical sentences (see Chap. 8, Sect. 8.4)—that is, analytic, according to Carnap’s intentions—but still unprovable. Carnap formulated the full definition of analyticity in Language II and also achieved some remarkable results, namely the fixed-point theorem, a limited version of the truth-undefinability theorem, and the definition of mathematical truth by evaluations (see Chap. 8, Sect. 8.5 for these results in the **STT** setting), which are just semantic from the contemporary point of view. Yet all of that was considered by Carnap as syntactic or quasi-syntactic. He remarked (Carnap 1934, p. 216):

> […] *truth and falsehood are not proper syntactical properties*; whether a sentence is true or false cannot generally be seen by its design, that is to say, by the kinds and serial order of its symbols. […] This fact has usually been overlooked by logicians, because, for the most part, they have been dealing not with descriptive but only with logical languages, and in relation to these, certainly, ‘true’ and ‘false’ coincide with ‘analytic’ and ‘contradictory’, respectively, and are thus syntactic terms. […].

Already around 1933, Carnap had in his hands all basic devices needed to build a formal semantics about 1933. In particular, Language II was sufficiently strong to embed in it semantic concepts. Yet he insisted that he worked with quasi-syntax. This is a very good illustration of the prejudices mentioned by Gödel (see above).

Carnap’s conversion to semantics was the result not only of influences from Gödel and Tarski, but also internal difficulties within the syntactic metaphilosophical project. First of all, the collection of pseudo-object-sentences is given empirically (by examples), not by general criteria. Hence, it is not clear how to precisely delineate its actual scope. Secondly, Carnap’s understanding of logic was not quite correct. He contrasted pure logic and applied logic, and located logical analysis in the latter. However, the concept of applied logic has at least two different significations. First, we can consider applied logic as logical analysis and synthesis of switching circuits, for instance. This sense of applied logic is quite similar to applied mathematics—say in physics. Certainly, applied logic as the logic of science (it is the second meaning) is something different. It is important for any account of the correctness of the results of logical analysis. Its criteria of validity or decidability are of course different than in the case of pure logic. Even if we claim that pure logic consists of analytic sentences, this understanding of analyticity has no application in establishing the status of the statements that belong to the logic of science. Thus, Carnap's solution that the logic of science is scientific, because it consists merely of analytic assertions, was certainly over-simplified. There are also objections to the concrete examples given by Carnap, examples even simpler than those adduced above. For instance, Carnap says that the sentence (*) 'Babylon was treated in yesterday's lecture' is quasi-syntactical and can be transformed into (**) 'The word 'Babylon' occurred in yesterday's lecture'. Yet (*) and (**) are certainly not equivalent. One might lecture on Babylon without using the word 'Babylon', and in order to infer (*) from (**) one must assume that the word "Babylon" refers to Babylon, but it is impossible to know this without an interpretation of language.

To some extent Carnap tried to reconcile his former syntacticism with new ideas. To do that, he proposed that the concept of a quasi-syntactical sentence should be supplemented by a more general idea of quasi-logical sentences (Carnap 1942, p. 245/246):

> Many sentences in philosophy are such that, in their customary formulation, they seem to deal not with language but merely with certain features of things or events or nature in general, while a closer analysis shows that they are translatable into sentences of L-semantics [that is, purely logical semantics – J. W.]. Sentences of this kind might be called *quasi-logical* or cryptological. By translating quasi-logical sentences into L-terms, the philosophical problems involved will often become clearer and their treatment in terms of L-semantics more precise. The same problems can often also be formalized and then dealt with by syntactical methods if a suitable calculus corresponding to the semantical system in question and formalizing its L-concepts is constructed. This way of syntactical reformulation of philosophical problems has been dealt with [*Logical Syntax of Language*, Chap. 5]. The method of semantical formulation of philosophical problems is to be developed in an analogous way; it may sometimes turn out to be more appropriate than the syntactical method […]."

The advent of formal semantics largely deconstructed Carnap's syntactic metalogic and metaphilosophy—a fact he very quickly recognized. In his intellectual autobiography Carnap says (Carnap 1963, p. 60; recall also a passage quoted in the Introduction to this book):

> Even before the publication of Tarski's article [Tarski 1933 – J. W.] I had realized, chiefly in conversations with Tarski and Gödel, that there must be a mode, different from the syntactical one, in which to speak about language. Since it is obviously admissible about facts, and, on the other hand, Wittgenstein notwithstanding, about expressions of a language, it cannot be inadmissible to do both in the same metalanguage. […]. In the new metalanguage of semantics, it is possible to make statements about the relation of

designation and about truth. […]. 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.

Carnap, armed with semantics and its conceptual tools, reformulated his metaphilosophy by saying (Carnap 1942, p. 250):

[Philosophical sentences] may first be translated into semantical sentences and then, under suitable conditions, into syntactical sentences also.

Carnap and Tarski became symbols of the victory of semantic methods in philosophy and logic (see also Woleński 1999b). However, semantics as a philosophical device did not convince everybody, even in the Vienna Circle. Carnap described the situation years later (Carnap 1963, p. 61; see also Woleński 2018a):

When I met Tarski again in Vienna in the spring of 1935, I urged him to deliver a paper on semantics and on his definition of truth at the International Congress for Scientific Philosophy to be held in Paris in September. 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. But Tarski was very sceptical. He thought that most philosophers, even those working in modern logic, would be not only indifferent, but hostile to the explication of the concept of truth. I promised to emphasize the importance of semantics in my paper and in the discussion at the Congress, and he agreed to present the suggested paper.

At the Congress it became clear […] that Tarski's sceptical predictions had been right. To my surprise, there was vehement opposition even on the side of our philosophical friends. […]. Neurath believed that the semantical concept of truth could not be reconciled with a strictly empiricst and anti-metaphysical point of view. Similar objections were raised in later publications by Felix Kaufmann and Hans Reichenbach.

The orientation of contemporary logic, mathematical as well as philosophical logic, is decisively semantic. On the other hand, the process of reorientation from syntax to semantics took even longer than in the case logic than in analytic philosophy (see Mostowski 1966 for a summary of the development of mathematical logic and the foundations of mathematics in 1930–1964 with the particular stress on the role of semantic methods). Even textbooks of logic published in the early 1940s had a syntactic flavour. Quine 1940 can serve as an example. The last chapter (on metalogic) has 'Syntax' as its title. The word 'semantic' occurs only once (p. 24) and refers to "properties which arise from the meaning of the expressions" (designation, synonymity). In particular, semantic methods are not used in proving metalogical theorems. As late as in 1956 we can still read (Church 1956, p. 67):

In concluding this Introduction, let us observe that much of what we have been saying has been concerned with the relation between linguistic expressions and their meaning, and therefore belongs to semantics. […]. From time to time in the following chapters we shall interrupt the rigorous treatment of a logistic system in order to make an informal semantical aside.

Alonzo Church himself regarded semantics as philosophically very important and made many very essential contributions to it—in particular, to analysis of intensional contexts and naming-relation. Yet he distinguished rigorous exposition of a logistic (that is, formal) system and semantic aside being outside logic proper. The second could be only illustrative and informal. Mostowski 1948 seems to be the first textbook in which semantic methods are widely used in establishing metamathematical results; other examples are Scholz 1950 and Hermes 1952. This way of doing and teaching logic became standard in the 1960s and later, although sometimes we can also find exceptions (see words of Roger Lyndon quoted (**DG18 V**)).

May be that Church and Lyndon simply afraid to enter philosophical controversies around the concept of truth and falsehood, but today it is difficult to imagine that a textbook of logic be based exclusively on the syntactic paradigm or even limited to a discussion of logical values, would (Lyndon's phrase) "choose [...] two neutral objects 1 and 0 to serve respectively, instead, of truth and falsehood". An explanation why semantics entered logic relatively lately resides in the fact that this required the development of model theory as a separate field of mathematical logic. It was done by Abraham Robinson and Tarski in the 1940s and early 1950s. In the 1930s, Tarski himself did not attributed a major to logical significance to model theory (see Vaught 1986 for Tarski's works in model theory). That is documented by his research at that time, which was devoted to many topics—but not to the theory of models. When model theory became a fully legitimate field of mathematical logic and the foundations of mathematics, it promptly found its place in books and papers on logic. One point should be especially stressed. Semantic proofs are usually simpler that syntactic. That was known much earlier in the case of elementary logic. For example, a truth-table—that is, a semantic device—provide an easy way of checking whether formulas are tautologies or not, at least in these cases where an expression does not contain too many variables (in the era of computers this limitation is not very important). This procedure, however, is not available for undecidable theories—in particular, for first-order predicate calculus. Thus, method of finding counterexamples by interpretation is indispensable for establishing, that a formula is not a tautology of **FOL**. Note that the completeness property does not suffice for the parity of syntactic and semantic methods. The real force of semantics and its methods appears in metamathematics, where semantic proofs are much simpler and more intuitive than those performed with aid of purely syntactic means (see Chap. 8, Sect. 8.5). Moreover, semantics, at least in Tarski's version, tolerates non-constructive methods, and thereby goes beyond syntax.

6.6 Formal Languages and Formal Semantics

There are currently two uses of the term 'formal semantics'. The first, becoming more and more popular among linguists, refers to a formal (mathematical) treatment of natural language (see Cann 1993, pp. 1–2) and its relation to the world. This approach accepts the following two theses: (a) natural language is a formal system (the Chomsky thesis); (b) there is no major difference between formal and natural language (the Montague thesis). Yet the real scope of the validity of (a) and (b) is still disputed. Hence, it is not clear to what extent speaking about the formal semantics of natural language is justified. Since I will not discuss this question at this point (see some additional remarks in Chap. 7, Sects. 7.2–7.5), let me immediately pass to the second notion of formal semantics, which refers to a rigorous mathematical treatment of formal languages and their semantic features in the extensional sense (recall that I do not consider intensional semantics), that is, related to truth and reference. Formal semantics in this sense is the same as model theory, where the model of a language is understood as defined in Chap. 5, Sect. 5.2.

Roughly speaking, to give a semantic characterization of a language **L** means to point out the class of models **Cl**(**M**) of **L**, that is, the class of structures which **L** is about. Apparently, everything is clear here. We have **L** as a formal language and **Cl** (**M**) as a formal set-theoretical object. Since both are mathematically characterized, one could say that everything appears as formal in formal semantics of formal languages. I will argue that this picture is essentially simplified, in particular, that not everything is formal in formal semantics, and what is more—namely that not everything could be.

Four different contrasts employed the study of languages are relevant to our problem (see also Chap. 7, Sect. 7.2 for a further relevant contrast):

(A) natural – artificial;
(B) informal – formal;
(C) unformalized – formalized;
(D) interpreted – non-interpreted.

For the first glance, it might seem that 'natural', 'informal', 'non-formalized', 'interpreted' express the same property, which can also be characterized as 'ordinary', 'colloquial' (Tarski used both adjectives), etc. Consequently, the words 'artificial', 'formal', 'formalized' and 'uninterpreted' also seem to refer to the same feature. However, closer inspection reveals that these identifications are quite dubious.

(**DG7**) My treatment of (A) and (B) is considerably influenced by Ryle 1953. In particular, the observation that the adjective 'ordinary' in the phrase 'the ordinary use of expressions' means something different than 'ordinary' in the phrase 'the use of ordinary expressions' is very relevant here. Gilbert Ryle points out that 'ordinary' in the first case can be replaced by 'standard'. Hence, the expressions of non-ordinary—for instance, mathematical—language have their standard use, although they do not belong to ordinary—that is colloquial—language. Of course, one should be very careful with declarations concerning the standard use of philosophical terms, but this issue, fundamental for metaphilosophy, must be left without further comments.►

If we say that a language is artificial, we usually have in mind that it was created for performing special tasks. Although an artificial language contains special symbols, it does not need to be formalized, for instance, Morse's alphabet is fully artificial, but not formalized. According to an elementary intuition, a language is just formal if its description proceeds independently of the content of its expressions and appeals only to their form. Finally, a formalized language is the result of a special process called formalization. Nothing precludes a formalized language from having been informal prior formalization. Note that the term ‘formalized’ refers to a result of act of formalization. For instance, formal mathematics arises by formalization of more or less informal mathematics. Furthermore, nothing precludes formalized languages from being interpreted (see Chap. 7 for more extensive remarks about this problem).

Contemporary mathematical linguistics succeeded in a very general description of formal languages (see Mateescu, Salomaa 1997a). Let **G** be an arbitrary non-empty set with an associative operation • and the neutral element **0** (which means that for any $g \in \mathbf{G}$, $g \bullet \mathbf{0} = g$). Thus, algebraically speaking, **G** is a monoid (a free semigroup). The elements of **G** are elementary strings (the alphabet), the operation • is interpreted as concatenation (in particular, we have that $g_i \bullet g_j = g_i\, g_j$) and **0** is the empty word. Every finite sequence of strings over the alphabet is counted as a word. Finally, a language **L** is a set of words over the (finite or infinite) alphabet **G**. No restriction is imposed as far as the nature of strings and words is concerned. Arbitrary objects (tables, chairs, English words, electric impulses, symbols of propositional calculus, etc.) can be employed as the building blocks of a language. Hence, languages understood as sets of words over alphabets are proper instantiations of the concept of formal language. In particular, they are fully compositional. Yet we encounter he following passage (Mateescu, Salomaa 1997a, p. 1):

> What is a language? By consulting a dictionary one finds, among others, the following explanations:
>
> 1. The body of words and systems for their use is common to people who are of the same community or nation, the same geographical area, or the same cultural tradition;
> 2. Any set or system of signs or symbols used in a more or less uniform fashion by a number of people who are thus enabled to communicate intelligibly with one other.
> 3. Any system of formalized symbols, signs, gestures, or the like, used or conceived as means of communicating thought, emotion, etc.
>
> The definitions 1–3 reflect a notion “language” general and neutral enough for our purposes.
>
> Further explanations are more closely associated with the spoken language […]. When speaking of formal languages, we want to construct formal grammars for defining language rather than to consider a language as a body of words somehow given to us or common to a group of people. Indeed, we will view a *language* as a set of finite strings of symbols from a finite alphabet. Formal *grammars* will be devices for defining specific languages. Depending on the context, the finite strings constituting a language can be also referred to as *words*, *sentences*, *programs*, etc. Such a formal idea of a language is compatible with the definitions 1–3, although it neglects all semantic issues and is restricted to *written* languages.
>
> The idea of a *formal language* being a set of finite strings of symbols from a finite alphabet constitutes the core of this Handbook. Certainly all written languages, the natural, programming or any other kind, are contained in this idea.

Although Mateescu and Salomaa say that the definition of a formal language "neglects all semantic issues", their approach is clearly intended to serve as a formal representation of languages, understood particularly as sets of sentences. This fact connects the description of formal language occurring in the last quoted passage with the account of language employed in metalogic.

As I have already noted, formal semantics proceeds by a rigorous (mathematical) treatment of relations holding between languages and the world. The semantics for **PC** and **FOL** as outlined in Chap. 5, Sects. 5.2.2, 5.2.4 provides good examples in this respect. In particular, it is easy to show that $\mathbf{L_{PC}}$ and $\mathbf{L_{FOL}}$ perfectly satisfy the algebraic definition of language, although logicians proceed in a less abstract way and prefer analysis in terms of alphabet, formation rules, and well-formed formulas, particularly sentences. The world of formal semantics can be represented by a mathematical structure (or rather the class of such structures). Let me recall that there must be a correlation between the signature of **L** and the signature of **M**. Thus, it is not quite correct to say that a purely formal language (at least, as defined as above) does not yield any information about the structure of its possible models, because if we know the signature of a language, we can infer from it some properties of models—for example, the number of distinctive objects and the arities of attributes (properties) and relations. The converse dependence (due to the fact that the interpretation of any language operates by functional dependencies) is even stronger. In particular, if the signature of a model (or interpretation) is <0*, 0*; 1*; 1*, 2*>, **L** must have two individual constants, one function symbols, and two predicate letters of which one is monadic and the other is dyadic. Thus, information to be derived from **L** about the structure of a model appears weaker, than the message accessible from the reverse direction. Incidentally, this circumstance falsifies an argument against analytic philosophy, namely that logical analysis replaces the world by language used in its description and thereby convert ontology into idealism.

The foregoing analysis shows that **L**, **Cl**(**M**), as well as the valuation function (as defined in the previous chapter), based on the concept of satisfaction (or truth) are necessary components of the formal semantics of a formal (or formalized) language **L**. In particular, the valuation function provides a correlation between expressions and appropriate items from models. What is formal in the description of **L**? The answer depends on how the adjective 'formal' is understood. Suppose that something qualifies as formal when it is independent of the content and dependent solely of the form. This antiquated formulation is useful, even if not quite precise. The matter becomes simpler in the case languages defined algebraically, because its structure and signature are given purely syntactically. The meaning (sense) of expressions of **L** is not relevant here. Since we assume a signature parallelism of **L** with its possible models, this correspondence determines the formal structure of both relata of the global semantic relation. On the other, the valuation function itself is not formal, because it depends of how expressions of **L** are understood. Take once more the example considered in (**DG9V**) and recall that the interpretation, which is historically faithful considers the sentence 'Socrates is a teacher of Plato' as true, but the sentence 'Aristocles was the teacher of Socrates' as false, although different understanding of the related constants cannot be excluded on purely semantic grounds. Yet, independently of whether particular interpretations function faithfully vis-a-vis empirical data or meanings present in ordinary languages, or whether the opposite holds, the valuation function operates in accordance with some definite material contents. Although we can still say that the content of an interpretation is formally given to some extent, because its description employs various mathematical devices, like functions, sets, relations, etc., the meaning of 'formal' is different here than when we speak about formal language as a free semigroup. Regardless, one is entitled to say that the semantic content of an interpretation is given informally as compared with the syntactic form of a language. This recalls the relation between formal and informal mathematics. We need the latter in order to speak about the former, although informal mathematics is also in some degree formal.

A general framework for the relation between the form and content of an interpretation can be outlined in the following way. Even in the case of syntax, syntactic description of a formal language **L** requires a partially informal metalanguage **ML** (this distinction will be elaborated in Chap. 7, Sect. 7.4). For example, the treatment of languages as sets of sentences clearly displays some informal intuitions. This situation appears much more explicit in the case of semantics. Evert Beth once remarked (Beth 1962, p. XIV):

> By *semantics* I mean a rigorously deductive treatment of the connections between the logical and mathematical symbols and the objects which they denote. An informal discussion of the same subject is denoted as *hermeneutics*.

Beth, as it follows clearly from his subsequent remarks (see, for example, p. 67, p. 72), he is inclined to consider hermeneutics (in his sense) as an informal discussion that fulfills ordinary (mathematical) intuitions. My proposal sharpens this view in two following directions. I would like to claim: (a) some hermeneutics governs any decision on how to correlate **L** and its interpretation; (b) some hermeneutics governs how to articulate (this point will be made more precise in Chap. 7, Sect. 7.4) expressions of **L** in **ML**. Both claims suggest that it is impossible to perform investigations in formal semantics without a certain amount of hermeneutics or, equivalently, that formal semantics is always embedded in a hermeneutics. This point can be metaphorically expressed by saying that the content of an interpretation goes from top to bottom, that is, from **ML** to **L**—but not conversely. In the beginning was (and is) the *interpreted* word, not the word itself. This maxim, as I will show in the next chapter, was central for Tarski (I label it as Tarski's central maxim).

The above discussion illuminates the issue, which I discussed in the previous section, of how to speak correctly about the totality of types. It is clear that we have to make some statements that refer to all types or all languages, although the stock of related assertions has to be limited in order to avoid inconsistencies. Now if we agree that formal semantics assumes some hermeneutics, the following proposal seems to be reasonable. For example, the theory of logical types, or the theory of syntactic categories function as formal constructions fully applicable only to families of specifically defined languages—formal or formalized. The proper way of explaining of these formalisms work has to informal, even in the case of very sophisticated methods such as arithmetization. Generally speaking, the fundamental results of metamathematics suggest that (a) semantics is prior to syntax (see more in Chap. 8, Sects. 8.4–8.6); and (b) what is informal is conceptually prior to what is formal. Now the rules of formal syntax and semantics that are applicable to formal constructions, have a limited importance in their informal expositions, even if the latter have a partially mathematical tenor. Of course, special care is call to avoid antinomies or unclarities, but the use of informal parlance appears indispensable to performing formalizations. Anticipating remarks in Chap. 8, Sect. 8.2, this situation recurs on the level of metasemantics, when metalanguage is formalized. Thus, the relation between the formal and the informal in semantics recalls the problem of the degrees of logicality (see Chap. 5, Sect. 5.5).

I would like to illustrate the last point by two important historical examples. Husserl (see Husserl 1929, Sects. 14–15) distinguished consequence-logic (the logic of non-contradiction) and truth-logic. He writes (pp. 54–55):

> […] questions concerning consequence and inconsistency can be asked about judgments *in forma* without involving the least inquiry into truth or falsity and therefore without ever bringing the concepts of truth and falsity, or their derivatives into the *theme*. […] we distinguish a level of formal logic that we call *consequence-logic* or *logic of non-contradiction*. […]. The fundamental concepts of pure analytics [= consequence-logic – J. W.] in the pregnant sense include, *as fundamental concepts of validity* […] *only analytic consequence and analytic contradiction*; as already said, *truth and falsity* […] are *not present* among them. This must be rightly understood: They are not present as fundamental concepts pertaining to the *thematic* sphere. Therefore, in this pure analytics, they play only the role that is theirs in all the science, so far as all sciences strive for truths and consequently talk about truth and falsity; but that is not to say that truth and falsity belong among the "fundamental concepts" of every science, the concepts pertaining to the proper essence of its particular scientific *province*. […]
>
> Inquiry for formal laws of possible *truth* and its modalities would be of a higher logical inquiry, *after* the isolation of pure analytics. If a logic restricts itself to the bare forms of the significations of statements – that is, the judgement-forms – what means does it have of becoming a genuine logic of truth? One can see forthwith that *non-contradiction* is an essential condition for possible *truth*, but also that mere analytics becomes converted into a *formal truth-logic* only by virtue of a *connexion* between these intrinsically separable concepts, a connexion that determines an eidetic law and, in logic, *must be formulated separately*.

I will not enter into an exegesis of Husserl's views on logic. Suffice to note that the logic of non-contradiction is syntactic in nature, whereas truth-logic plays the role of semantics. Further, Husserl claims that pure analytics should be supplemented by a formal truth-logic, because the latter goes beyond the former. These views are fairly close to what contemporary logicians say about the relation between syntax and semantics. On the other hand, it is unclear whether Husserl applies the adjective 'formal' in the same sense when he speaks about pure formal analytics and when he speaks about the formal logic of truth. As I tried to show, these two issues should be sharply distinguished. Perhaps a more important question revolves around Husserl's view pertaining to the justification of logic. He maintained that an eidetic analysis would be necessary for this task. Thus, we have three different levels: (a) pure formal analytics (syntax); (b) formal logic of truth (formal semantics); (c) eidetics (transcendental logic). According to Husserl, analysis on the level (c) is the strongest and most reliable from the methodological point of view than that available on levels (a) and (b). However, even if one claims that logic requires a justification by appealing to its philosophical (in my scheme, semantic) foundations, it cannot be stronger than available on the levels (a) and (b). It seems to me that it is a mistake on the part of all conceptions of logic which consider transcendental logic as prior to formal logic. In the case of Husserl this seems to be derivative of his belief that philosophy should be presuppositionless.

The second example concerns frequent interpretations of formalism in the philosophy of mathematics. It is said that Hilbert required the complete formalization of the whole of mathematics, including the way of explaining what formalization looks like. Yet this is a misunderstanding of Hilbert's original position, because he distinguished formal and intuitive mathematics and claimed that the latter should be formalized by the standards envisaged for the former. For Hilbert, finitary methods are the most secure basis for mathematics, for instance, in checking consistency. On the other hand, Hilbert knew perfectly well that to some extent his program of formalization has to be informally explained and carried out. Various erroneous comments on formalism ignore the fact that talk about reliable proof-techniques, even when embedded in purely logical vocabulary, assumes a prior informal way of speaking Historically, it is rather interesting that Husserl and Hilbert—who disagree almost on every issue—could agree on how the formal is related to the informal.

(**DG8**) Below is a nice quotation illustrating the relation between the formal and informal. The quotation is taken from a specialized monograph on computable set theory (Canone, Omodeo, Policriti 2001, p. 15):

> A formalism is not simply a language, a language is not merely a collection of expressions. In natural language an expression has a meaning, and it is generally desirable that such meaning be unique. What can we say about formal languages, which in addition to being artificial are designed with the aim of separating from the content? Separating form from content does not mean abandoning altogether any concern about the meaning of expressions, but rather leaving open a large variety of interpretations for the same expression. Investigating this variety of meanings is the task of *semantics*.

This view perfectly concurs with Tarski's central maxim.►

References

Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E., Romanowska, A., Woleński, J. (Eds.) (2004). *Provinces of Logic Determined. Essays in the Memory of Alfred Tarski. Parts I, II and III* (*Annals of Pure and Applied Logic*, 127). Amsterdam: Elsevier.

Ajdukiewicz, K. (1993). Kategorie syntaktyczne i antynomie logiczne. Wykłady z semantyki logicznej w r. a. 1930/1931 (Syntactic Categories and Logical Antinomies. Lectures on Logical Semantics in the Academic Year 1930/1931). *Filozofia Nauki* I/1, 163–182.

Allan, K. (Ed.). (2009). *The Concise Encyclopedia of Semantics*. Amsterdam: Elsevier.

Bar-Hillel, Y. (1970). *Aspects of Language*. Amsterdam: North–Holland.
Bar-Hillel, Y. (1970a). Husserl's conception of a purely logical grammar. In Bar-Hillel, 1970, 89–97.
Beth, E. W. (1962). *Formal Methods. An Introduction to Symbolic Logic and to the Study of Effective Operations*. Dordrecht: Reidel.
Cann, R. (1993). *Formal Semantics. An Introduction*. Cambridge: Cambridge University Press.
Cantone, D. Omodeo, E., Policriti, D. (2001). *Set Theory for Computing: From Decision Procedures to Declarative Programming with Sets*. Berlin: Springer.
Carnap, R. (1934). *Logische Syntax der Sprache*. Wien: Springer; Engl. tr., London: Routledge and Kegan Paul 1937.
Carnap, R. (1934a). Die Antinomien und die Unvollständingkeit der Mathematik. *Monatshefte für Mathematik und Physik,* 41, 263–284.
Carnap, R. (1939). *Foundations of Logic and Mathematics*. Chicago: University of Chicago Press.
Carnap, R. (1942). *Introduction to Semantic*. Cambridge, Mass: Harvard University Press; 2nd ed., 1958.
Carnap, R. (1963). Intellectual autobiography. In Schilpp (1963), 3–84.
Chappell, V. C. (Ed.) (1964). *Ordinary Language*. Englewood-Clifts: Prentice-Hall.
Church, A. (1956). *Introduction to Mathematical Logic*. Princeton: Princeton University Press.
Coffa, A. (1987). Carnap, Tarski and search for truth. *Nous,* 21, 547–572.
Coffa, A. (1991). *The Semantic Tradition from Kant to Carnap. To the Vienna Station*. Cambridge: Cambridge University Press.
Copi, I. (1971). *The Theory of Logical Types*. London: Routledge and Kegan Paul.
Edwards, P. (1967). *The Encyclopedia of Philosophy*, v. 1–8. New York: Macmillan.
Feferman, S. (1984). Kurt Gödel: conviction and caution. *Philosophia Naturalis*, 21, 546–562; repr. in Feferman (1998), 150–164.
Feferman, S. (1998). *In the Light of Logic*. Oxford: Oxford University Press.
Fraenkel, A., Bar-Hillel, Y., Levy, A. (1973). *Foundations of Set Theory*. Amsterdam: North-Holland.
Gödel, K. (1930). Die Vollständigkeit der Axiome des logischen Funktionenkalkül. *Monatshefte für Mathematik und Physik*, 37 349–360; repr. with Eng. tr. in Gödel (1986), 102–123.
Gödel, K. (1931). Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38, 173–198; repr. with Eng. tr. in Gödel (1986), 145–195.
Gödel, K. (1986). *Collected Works*. Oxford: Oxford University Press.
Hermes, H. (1952). *Mathematische Logik*. Leipzig: Teubner.
Henkin, L., Addison, J., Chang C. C., Craig, W., Scott, D., Vaught, R. (Eds.) (1974). *Proceedings of the Tarski Symposium. An International Symposium Held to Honor Alfred Tarski on the Occasion of His Seventieth Birthday*. Providence, R. I.: American Mathematical Society.
Hintikka, J. (1997). *Lingua Universalis vs. Calculus Ratiocinator: An Ultimate Presupposition of Twentieth-Century Philosophy*. Dordrecht: Kluwer.
Hipkiss, R. A. (1995). *Semantics. Defining the Discipline*. Mahwah: Erlbaum.
Hiż, H. (2004). Reexamination of Tarski's semantics. In Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E., Romanowska, A., Woleński, J. (2004), 39–48.
Husserl, E. (1929). *Formale und transzendentale Logik. Versuch einer Kritik der logischen Vernuft*. Halle: Niemeyer; Eng. tr. by D. Cairns, The Hague: Nijhoff 1969.
Husserl, E. (1984). *Einleitung in der Logik und Erkenntistheorie (Vortrag 1906–1907).* The Hague: Nijhoff; Eng. tr., Dordrecht: Springer 2008.
Kotarbiński, T. (1929). *Elementy teorii poznania, logiki formalnej i metodologii nauk* (Elements of Theory of Knowledge, Formal Logic and Methodology of Science). Lwów: Ossolineum; Eng. tr., *Gniosiology. The Scientific Approach to the Theory of Knowledge.* Oxford: Pergamon Press.
Kreisel, G. (1987). Gödel's excursions into intuionistic logic. In Weingartner, Schmetterer (1987), 65–186.
Kusch, M. (1989). *Language as Calculus vs. Language as Universal Medium. A Study in Husserl, Heidegger and Gadamer*. Dordrecht: Kluwer.

Leśniewski, S. (1929). Grundzüge eines neuen System der Grundlagen der Mathematik. *Fundamenta Mathematicae*, 14, 1–81; Eng. tr. (by M. P. O'Neil) in Leśniewski (1992), 410–605.

Leśniewski, S. (1992). *Collected Works*. Dordrecht: Kluwer.

Mateescu, A., Salomaa, A. (1997a). Formal languages: An introduction and a synopsis. In Rosenberg, Salomaa (1997), 1–39.

Morris, Ch. (1938). *Foundations of the Theory of Signs*. Chicago: University of Chicago Press.

Mostowski, A. (1948). *Logika matematyczna* (Mathematical Logic). Warszawa: Monografie Matematyczne.

Mostowski, A. (1967). Alfred Tarski. In Edwards (1967), v. 8, 77–81.

Mycielski, J. (2004). On the tension between Tarski's nominalism and model theory (Definitions for a Mathemathical Model of Knowledge). In Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E., Romanowska, A., Woleński, J. (2004), 215–224.

Ogden, C. K., Richards, I. A. (1923). *The Meaning of Meaning: A study of the Influence of Language upon Thought and of the Science of Symbolism*. London: Routledge and Kegan Paul.

Quine, W. V. O. (1940). *Mathematical logic.* Cambridge, Mass.: Harvard University Press.

Quine, V. V. O. (1953). From a Logical point of View. Cambridge, Mass.: Harvard University Press.

Quine, V. V. O. (1953a). Notes on the theory of reference. In Quine (1953), 130–138.

Ramsey, F. P. (1925). The foundations of mathematics. *Proceedings of the London Mathematical Society* (Series 2), 25, 338–384; repr. in Ramsey (1931), 1–61 and Ramsey (1978), 152–212.

Ramsey, F. P. (1931). *The Foundations of Mathematics*. London: Routledge and Kegan Paul; 2nd ed., *Foundations. Essays in Philosophy, Logic, Mathematics and Economics*. London: Routledge and Kegan Paul.

Rosenberg, G., & Salomaa, A. (Eds.). *Handbook of Formal Languages, v. 1: Word, Language, Grammar*. Berlin: Springer.

Russell, B. (1903). *The Principles of Mathematics*. Cambridge: Cambridge University Press. Cambridge.

Ryle, G. (1953). Ordinary language. *Philosophical Review* , XLII, 167–186; repr. in Chappell (1964), 24–40.

Scholz, H. (1950). *Vorlesungen über Grundzüge der mathematischen Logik*. Münster: Aschendoffsche Verlagsbuchhandlung.

Schilpp, P. A. (Ed.) (1963). *The Philosophy of Rudolf Carnap*, 3–84. La Salle: Open Court.

Sierpiński, W. (1965). *Cardinal and Ordinal Numbers*. Warszawa: Państwowe Wydawnictwo Naukowe.

Tarski, A. (1930–1931). O pojęciu prawdy w odniesieniu do sformalizowanych nauk dedukcyjnych (On the Concept of Truth in Reference to Formalized Deductive Sciences). *Ruch Filozoficzny*, 12, 210–211; repr. in Tarski (1986), v. 4, 559.

Tarski A. (1931). Sur les ensembles définissables de nombres réels. I. *Fundamenta Mathematicae*, XVII, ss. 210–239; Eng. tr. in A. Tarski, 1956, 110–142.

Tarski, A. (1932). Der Wahrheitsbegriff in der Sprachen der deduktiven Wissenschaften. *Akademie der Wissenschaften in Wien. Matematisch-Wissenschaftliche Anzeiger*, 69, 23–25; repr. In Tarski (1986), v. 1, 613–617.

Tarski, A. (1933). *Pojęcie prawdy w językach nauk dedukcyjnych* (The Concept of Truth in Languages of Deductive Sciences). Warszawa: Towarzystwo Naukowe Warszawskie; Germ. tr. (with additions) as Tarski (1935).

Tarski, A. (1935). Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica*, 1, 261–405; repr. in Tarski (1986), v. 2, 51–198. Engl. tr. (The Concept of Truth in Formalized Languages), in Tarski (1956), 152–278 (page-references to Tarski (1933).

Tarski A. (1936). Grundlegung der wissenschaftliche Semantik. In *Actes du Congès International de Philosophie Scientifique*, v. 3), Paris: Herman, 1–8; Eng. tr. in A. Tarski (1956), 401–408.

Tarski, A. (1954). [Contribution to the Discussion of P. Bernays, Zur Beurteilung der Situation in der beweistheoretischen Forschung]. *Revue Internationale de Philosophie*, 8, 16–20; repr. in Tarski (1986), v. 4, 713–714.

Tarski, A. (1954a). [Remarks of Alfred Tarski]. *Revue Internationale de Philosophie*, 27–28, 18–19; repr. in Tarski (1986), v. 4, 713).

Tarski, A. (1956). *Logic, Semantics, Metamathematics*. *Papers of 1923 to 1938* (tr. by J. H. Woodger). Oxford: Clarendon Press; 2nd ed., Indianapolis: Hackett Publishing Company.

Tarski, A. (1986). *Collected Papers*, v. 1–4. Basel: Birkhäuser, 2nd ed., 2018.

Tarski, A. (1986a). What are logical notions? *History and Philosophy of Logic,* 7, 143–154.

Tuboly, A. T. (2017). From 'Syntax to 'Semanik'—Carnap's Inferentialism and Its Prospects (unupblished manuscript).

Van Heijenoort, J. (1985). *Selected Essays*. Napoli: Bibliopolis.

Vaught, R. (1974). Model theory before 1945. In Henkin, Addison, Chang, Craig, Scott, Vaught (Eds.) (1974), 153–172.

Vaught, R. (1986). Alfred Tarski's work in model theory. *Journal of Symbolic Logic*, 51, 861–882.

Weingartner, P., & Schmetterer, L. (Eds.) (1987). *Gödel Remembered*. Napoli: Bibliopolis.

Whitehead, A. N., Russell, B. (1910–1913). *Principia Mathematica.* Cambridge: Cambridge University Press; 2nd ed., 1925.

Woleński, J. (1995a). On Tarski's background. In J. Hintikka (Ed.). (1995), 331–341); repr. in Woleński (1999), 126–133.

Woleński, J. (1999). *Essays in the History of Logic and Logical Philosophy*. Kraków: Jagiellonian University Press.

Woleński, J. (1999b). The semantic revolution—Rudolf Carnap, Kurt Gödel, Alfred Tarski. In Woleński, Köhler (1998), 1–15.

Woleński, J. (2018a). The semantics controversy at the Paris Congress 1935. *Philosophia Scientiæ*, 22(3), 3–15.

Woleński, J., Köhler, E. (Eds.) (1999). *Alfred Tarski and the Vienna Circle*, Kluwer: Dordrecht.

Chapter 7
Semantic Theory of Truth—Informal Aspects

Abstract This chapter outlines an intuitive approach to **STT**, very closely related to Tarski's original approach. The essential role of the Liar antinomy and its solution by introducing the language/metalanguage distinction is stressed as well as the role of interpreted languages is pointed out. Finally, heuristics of forming the semantic truth definition via the concept of satisfaction is reported.

7.1 Introduction

Let me recall (see Introduction) that **STT** is considered in this book in the frameworks of **FOL**. Clearly, this choice determines which formal tools are proper for shaping this theory as a chapter of formal semantics, metalogic or metamathematics (see (**SDfVER**) in Chap. 5 as a temporary illustration; the full formal definition of truth will be provided in Chap. 8, Sect. 8.2). However, Tarski also explained the proposed truth-definition in an informal way and regarded this manner of introducing related ideas as very important (see Woleński 2016b). This chapter focuses on this last approach and tries to exhibit those intuitions that motivated formalism. I will not delve very much in philosophical issues, because these will discussed in Chap. 9, but rather concentrate on intuitions related to informal semantics. As I already said (see Chap. 5, Sect. 5.1), the original version of **STT** was framed in a formalism similar to the simple theory of types (in fact, modified by the theory of semantic categories—see Chap. 5, Sect. 5.3). Fortunately, Tarski's intuitive background is largely invariant with respect to choosing the theory of types or **FOL** as formal foundations of **STT**. The present chapter consists of four sections. Section 7.1 briefly discusses various meanings of the adjective 'informal' as contrasted with 'formal'. In a sense it is a supplement to Chap. 6, Sect. 6.6. The Liar Paradox (**LP**) is the subject of Sect. 7.2. The next section pertains to the resolution of **LP** in **STT** via the language/metalanguage distinction as well as various problems related to this pair of concepts. Finally, Sect. 7.4 is devoted to the conditions of adequacy for **STT** as formulated by Tarski himself and, reconstructs his heuristics.

(**DG1**) There is no sufficient historical evidence for answering the question 'Why did Tarski become interested in the problem of truth?". Of course, it was perhaps one of the most celebrated philosophical problems in Polish philosophy after Twardowski, and all of Tarski's principal teachers, Leśniewski, Łukasiewicz and

J. Woleński, *Semantics and Truth*, Logic, Epistemology, and the Unity of Science 45, https://doi.org/10.1007/978-3-030-24536-8_7

Kotarbiński (to mention these names once again), were strongly involved in various investigations related to the concept of truth. On the other hand, in the 1920s Tarski concentrated on advanced logic and fairly specialized foundations of mathematics, including abstract set theory, topology, geometry and arithmetic of real numbers, that is, topics rather remote from philosophy. According to Tarski himself, his interest in the concept of truth was sparked in 1929 (see Tarski 1933, p. 154). In 1930, he delivered three lectures, one (probably in December) before of the Polish Mathematical Society, The Lvov Branch (Tarski 1930, see also Tarski 1931), a second (on October 8) at the Warsaw Philosophical Institute, and third (on December 15) before the Polish Philosophical Society in Lvov (see Tarski 1930–1931; it is an abstract based on the two previous). The first concerned definable sets of numbers. Tarski observed that the concept of satisfaction is essential for definability. The next two lectures outlined Tarski's truth-definition. Twardowski noted (see Twardowski 1997, v. 2, p. 180):

> Monday, December 15 […]. In the evening, Tarski's lecture in P. T. F. [Polish Philosophical Society] – very interesting as well as very beautifully constructed.

Tarski lectured on truth in Vienna on January 21, 1932 (see Tarski 1932). The content of this talk followed his lectures on truth in Warsaw and Lvov, but with some additions concerning the undefinability of truth (see Chap. 8, Sect. 8.5; I return to some historical questions in this fragment of the book).

Tarski's main essay on the concept of truth (Tarski 1933) was communicated to the Warsaw Scientific Society in 1931, but published two years later. We do not know whether the published version was the same as that submitted in 1931 (Jerzy Łoś told me once that Tarski postponed the publication, because he wanted to rethink Gódel's result concerning the incompleteness of arithmetic, but no evidence confirms this suggestion). On the other hand, comparing Tarski's lectures on truth in 1930 with the main monograph, one can see, that the concept of satisfaction, so essential for **STT** (see Sect. 7.3), was directly motivated by his mathematical investigations on definability of sets of real numbers and reported in one of lectures in Lvov (see above). The concept of satisfaction was loosely used by mathematicians for a long time, but Tarski, guided by needs of metamathematics, decided to define it in a precise manner.

The German translation of Tarski 1933 appeared in 1935 (see Tarski 1935). It contains essential changes (see Gruber 2016; Woleński 2017a for details) in comparison to the Polish original. Leopold Blaustein, a Polish philosopher (but not a logician), was the main translator, but Ajdukiewicz, Carnap, Maria Kokoszyńska, Popper and Twardowski were also involved in the process of translation to greater or lesser extent. This is a significant historical event, because it provides additional evidence that Tarski's ideas were known to Ajdukiewicz, Carnap, Kokoszyńska and Popper before the German translation of Tarski 1933 was published (moreover, as it was customary at that time, offprints of Tarski's essay were distributed earlier). All four philosophers became very soon defenders of semantics against it critics (see Chap. 6, Sect. 6.5).

In 1935, Tarski delivered two lectures in Paris at the International Congress of Scientific Philosophy (see Tarski 1936, Tarski 1936a); both contain various references to **STT**. Tarski 1944, a very important paper, discusses **STT** from a formal as well as philosophical point of view, as well as affords rejoinders, mostly philosophical, to various critical remarks addressed against semantics and its uses in philosophy. The book Tarski, Mostowski, Robinson 1954 contains important treatment of the undefinability of truth (see more in Chap. 8, Sect. 8.8). As I already indicated in Chap. 5, Sect. 5.1, the paper Tarski, Vaught 1957 introduced the concept of truth for first-order languages, but this work does not discuss intuitive issues. Finally, Tarski 1969 concludes Tarski's writings that have a direct bearing on the concept of truth. His various shorter remarks related to this concept will be mentioned on suitable places. Basically, the above survey omits, except one remark on the translation of Tarski 1933, the international context of the reception of **STT** (see Introduction for this issue).►

7.2 STT, the 'Formal' and Its Opposites

STT has several ingredients (n particular, it is not proper to reduce this theory to the definition proposed by Tarski; this simplification is frequently made by critics of Tarski—Black 1948 is an example). The most important (at least according to the analysis in this book) are as follows (see also Stegmüller 1957, De Florio 2013):

(I) Truth as a property of sentences (see Chap. 4, Sect. 4.2)
(II) Relations between truth and meaning;
(III) Diagnosis of semantic paradoxes;
(IV) Resolution of semantic paradoxes;
(V) Relativization to languages;
(VI) **T**-scheme (*A* is true if and only if *A*);
(VII) The principle (**BI**) (of bivalence);
(VIII) Material and formal adequacy of truth-definition;
(IX) Conditions imposed on a metalanguage in order to obtain a proper truth definition;
(X) The relation between language and metalanguage;
(XI) The truth-definition itself;
(XII) Maximality of the set of truths in a given language;
(XI) The undefinability theorem.

In fact, the above points provide a general synopsis of the considerations in my subsequent analysis in this chapter as well as in Chaps. 8 and 9, but at this juncture they are mainly mostly intended to help identify what is formal and what is informal in **STT**. This issue is crucial for interpreting **STT**, due to the fact that many commentators (see references in various places of this book) argues that this theory applies to formal languages only. I will argue that this view is improper.

Here is the list of meanings of the adjective ‘formal’ and the terms contrasted to it (that is, falling under the general rubric indicated by the adjectives ‘informal’ and ‘non-formal’), listed in entries in various philosophical dictionaries and encyclopedias (particular items in this list are not mutually exclusive and some of them overlap; moreover, I only take into account these meanings which can be presumably relevant for analysis of **STT**):

(A) Formal versus substantive (form versus content);
(B) Formal versus material;
(C) Formal as computable versus non-computable;
(D) Formal (effective, finitary, etc.) versus non-finitary;
(E) Formal as logical versus extralogical;
(F) Formal as syntactic versus semantic;
(G) Formal as extensional vs intensional;
(H) Formal as compositional versus non-compositional;
(I) Formal as logical plus mathematical versus empirical;
(J) Formal as analytic versus synthetic;
(K) Formal as in ‘formal language’ versus ordinary (informal) language.

Some comments are in order about (A)–(K) in the context of **STT**.

Ad (A) Tarski claimed that **STT** explains the content of the concept of truth. If so, this theory is substantive. On the other hand, **STT** is formal in the sense that logical apparatus plays a crucial role in the analysis of this concept. Obviously, there is no contradiction between both claims (i) **STT** is substantive; (ii) **STT** is formal.

Ad (B) Epistemology traditionally distinguishes so-called formal truths (truths of logic and mathematics) and material truths concerning, roughly speaking, how things are. **STT** is a unified theory of both kinds of truth.

Ad (C)–(D) These issues pertain more to the formal aspects of **STT** and will be considered in Chap. 8.

Ad (E) If various logical tools are used essentially in **STT**, the question arises whether this theory as such is logical or extralogical. We can also ask which particular elements of **STT** are logical or extralogical, for instance, whether **T**-scheme is a logical tautology or not. Any answer heavily depends on the view on the scope of logic. Recall (see Chap. 6, Sect. 6.3) that Tarski himself was skeptical about the possibility of a precise borderline between logic and, so to speak, extralogic. If we accept **FOT** (see Chap. 5), **STT** is formal, but extralogical, but if the domain of logic is extended, the theory in question is logical.

Ad (F)–(H) Considerations in Chap. 6, Sects. 6.4–6.6 concluded that **STT** is semantic, extensional and compositional. However, some authors claim that it is actually syntactic and that extensionality and/or compositionality impose essential limitations on it. See Chaps. 7 and 8 for further comments on this topic.

Ad (I) As in Ad (B), **STT** is applicable to all kinds of truth—logical, mathematical as well as empirical.

Ad (J) **STT** applies to analytic as well as synthetic truths.

Ad (K) Some critics of **STT** claim that this theory is not applicable to ordinary language, and argue that this circumstance constitutes its serious disadvantage. However, the following considerations show that this objection is not correct, although Tarski's views on this issue considerably evolved.

Comparing (I)–(XI) and (A)–(K) suggests that any analysis of **STT** should always check the interplay between its formal and informal aspects. As a good example one my take the case (see Chap. 5, Sect. 5.5) of the role of informal explanations in building pure logic, even by using partly informal metalogical devices (see Chap. 5). However, one should not expect that formal and informal aspects of **STT** (and any other construction based on logic sensu *stricto*) perfectly correspond. In fact, they are parallel to some extent only, because formalizations always introduce simplifications, idealizations, models, schematizations, etc., and these devices neglect various aspects of what becomes formalized. Thus, the results of formalizations appear as approximate pictures of the material being formalized. Incidentally, this circumstance very heavily influence most of philosophical discussions concerning results (in particular, what I call—interpretative consequences) supervening on employing logical devices in analysis performed by philosophers.

7.3 The Liar

As I noted in Chap. 6, Sect. 6.4, Tarski considered semantic antinomies as a very serious challenge for logic and semantics. Here is yet another documentation of his view (Tarski 1969, p. 409):

> Two diametrically opposed approaches to antinomies can be found in the literature of the subject. One approach is to disregard them, to treat them as sophistries produced mainly *pour épater de bourgeois*, as jokes which are not serious but malicious, and which at their best only give evidence of the cleverness of their authors. The opposite approach is characteristic of certain thinkers of the nineteenth century and is still programmatically represented, or was so a short while ago, in certain parts of our globe. According to this approach antinomies constitute a very essential element of human thought, they must appear again and again in intellectual activities, and their presence is the basic source of real progress. As often happens, the truth is probably somewhere in between. Personally, as a logician, I could not reconcile myself with antinomies as a permanent element of our system of knowledge. However, I am not the least inclined to treat antinomies lightly. The appearance of an antinomy is for me a symptom of a disease. Starting with premises which seem intuitively obvious, using forms of reasoning which seem intuitively certain, an antinomy leads us to a nonsense, a contradiction. Whenever this happens, we have to submit our ways of thinking to a thorough revision, to reject some premises in which we believed, or to improve some forms of argument which we used. We do this with the expectation that not only the old antinomy will be disposed of, but no new one will appear. To this end, we test our reformed system of thinking by all available means and, first of all, we attempt to reconstruct the old antinomy in the new setting (hoping, of course, that our attempts will fail); this testing is a very important activity in the realm of speculative thought, akin to carrying out crucial experiments in empirical science.

The quoted fragment outlines a program for dealing with antinomies. For Tarski, the danger of antinomies did not consist in the possibility of inferring any sentence from a contradiction via the theorem $A \wedge \neg A \Rightarrow B$ (*ex falso quodlibet*) of **PC**, but in the fact that, accepting inconsistency, we assert at least one false statement. This violates the principle that falsehoods are not to be accepted.

The Liar Paradox was Tarski's particular concern. Its history goes back to antiquity (see Rüstow 1908 for details); it is frequently called the Epimenides Paradox, due to Epimenides, the Cretean who invented it, according to tradition came out with it. The paradox is displayed by the sentence (the formulation appellation alludes to the fact that Creteans were famous in ancient Greece for lying)

(*) I am lying now.

Intuitively, if (*) is true, it is false (assuming that 'lying' means saying something falsely; in fact, this identification was obvious for thinking based on archaic Greek —see Chap. 1). On the other hand, if (*) is false, it is true. Contradiction! Now reflecting on the last quoted passage, a successful definition of the concept of truth cannot be achieved without resolving **LP**. This is just the old antinomy, but, according to Tarski, we should dress it in a more contemporary language and then try to test **LP** by modern logical standards. Although **LP** is interesting in itself a logical puzzle, it has a wider importance, just as a difficulty in truth-theory.

Polish logicians considered the formulation (*) as not quite satisfactory for its explicit temporal indexicality connected with the word 'now'. It was Łukasiewicz (see Łukasiewicz 1915) who formulated the version free of the mentioned defect (Tarski employed in his works mostly this version of **LP**; he also considered other formulations, but I omit the, because they do not add anything new). My presentation proceeds as follows (it uses a slightly modified reasoning of Tarski). Consider the sentence

(L) The sentence in the line indicated by the symbol (L) in this section is false.

Let l be an abbreviation for 'the sentence indicated by the symbol (L) in this section'. This convention empirically justifies

(1) l = the sentence l is false.

Using **T**-scheme, we obtain (identity justifies equivalence).

(2) l is true if and only if l.

Application of (1) gives

(3) l is true if and only if l is false.

Further, since

(4) l is false if and only if $\neg l$,

the contradiction

(5) $l \Leftrightarrow \neg l$,

follows. One could eventually say that the letter l plays the double role in (1) as a name inside the sentence 'the sentence l is false' and a sentential constant in before the sign =. However, this situation is created by self-reference involved in (1) consisting that the sentence l refers to itself.

(**DG2**) The mentioned temporal indexicality occurring in the original version of **LP** is eliminated in (L) by its formulation together with the convention (1). However, one can observe that the meaning of the sign = is not transparent. Clearly, this symbol does not refer to the identity predicate in the sense of **FOL**. It is rather a device similar to referring to participants in athletic competitions, for example, that a particular wears a jersey with, say, the number, 99 (like Wayne Gretzky one of the greatest stars in the entire history of NHL). This device allows for the identification of players in definite contexts—for instance, in the protocols of referees or the reports made by journalists. Similarly, in the above reasoning, (1) justifies to the substitute of right side for the left and conversely, but without suggesting that all properties of l and 'l is false' are the same. This last circumstance makes a difference between the symbol = in (1) and the identity predicate in **FOL**. In particular, the identity in (1) does not satisfy the Leibniz principle. We do not claim that l and 'the sentence l is false' share all properties, but only that introduced for the sake of argument. Note that the sentence 'All Creteans are lying' used by a Cretean does not lead to any paradox. If it is true, it is false, but it is false, implies that at least one Cretean said a true sentence.►

Tarski (see Tarski 1944, p. 672), following Leśniewski, argued:

> (I) We have implicitly assumed that the language in which the antinomy is constructed contains, in addition to its expressions, also the names of these expressions, as well as semantic term such as the term "true" referring to the sentences of this language; we have also assumed that all sentences which determine the adequate usage of this term can be asserted in the language. A language with these properties will be called "*semantically closed.*"
>
> (II) We have assumed that in this language the ordinary laws of logic hold.
>
> (III) We have assumed that we can formulate and assert in our language an empirical premise such as the statement (2) [(1) in my version—J. W.] which has occurred in our argument.

The third premise is not essential, but I will not go into its elimination. Perhaps it is sufficient to say that the presence of informal factors in arguing on formal issues justifies the use of (1) based on empirical findings. The first premise can be split into two parts: (a) the language, say **L**, contains expressions as well as their names (the semantic closeness), and (b) **T**-scheme holds universally (this scheme regulates the usage of the predicate 'is true'). The two parts of premise (I) (in the last quotation from Tarski) imply that semantic terms can be used self-referentially, that is, that the predicates 'is true' and 'is false' can refer to sentences in which they occur, like in the case of (L). To sum up, **LP** is generated by the following assumptions (since I discuss **STT**, I skip other diagnoses of **LP**):

(α) Self-referentiality (self-referring);
(β) **T**-scheme as universal, that is ‘for any *A*, *A* is true if and only if *A*’ (this equivalence will be analyzed below with respect to the universal quantification, so this version of **T**-scheme is temporary);
(γ) Classical (ordinary) logic.

Consequently, we have three different options for resolving **LP**. Firstly, we can eliminate self-referentiality; secondly, reject or modify **T**-scheme as a basic constraint for using the predicate ‘is true’, and, thirdly, change logic from classical to another system. Tarski choose the first alternative. It directly leads to modifying **T**-scheme to avoid the self-reference as a source of **LP** (and other semantic antinomies).

7.4 The Liar and the Language/Metalanguage Distinction

Assume that we use a language **L** for expressing our assertions about the world **W**. Although **L** belongs to (is a part of) **W**, most of our statements concern the extralinguistic reality, so to speak, in particular, things, facts, processes, states of affairs, etc. (I neglect that some philosophers are inclined to consider facts and states of affairs as linguistic items or related somehow to language. However, it also happens that we also make statements about **L**. That occurs, when we describe grammatical or logical properties of expressions, or simply quote phrases uttered by other people or our own past utterances. Here are examples:

(a) Every correct sentence consists of its grammatical subject and a verb;
(b) The sentence ‘Snow is white’ consists of three words;
(c) Socrates says in *Theaetetus*, 204e ‘Then the whole does not consist of parts. For if it did would be all the parts and so would be a sum’;
(d) Is the word ‘heterological’ heterological or autological?;
(e) ‘*A*’ is true if and only if f *A*;
(f) The sentence *l* is true;
(g) The sentence *l* is false.

There is a very simple difference between (a) on the one hand, and (b)–(e) on the other hand (I omit (f) and (g) for the time being, because they lead to special questions). The former expresses an elementary grammatical principle and does not refer to any concrete expression. In contradistinction, examples (b)–(e) mention some single words or complex expressions, the very ones that occurring between quotes.

Intuitively speaking, all the words in example (a) are used, but not mentioned. On the other hand, the quoted expressions in (b)–(e) are mentioned, but not used.

The first approximation of the language/metalanguage distinction points out that **L** uses expressions to speak about the world **W**, the metalanguage (**ML** for brevity) mentions some expressions of **L**, although it employs its own linguistic devices. If we confine ourselves to the simplest case, **L** is the object-language (referring to things in **W**), and **ML** is the metalanguage in the sense that it is about **L** and its particular items. In ordinary parlance, the distinction just adduced is vague, because the same language, English, German, Polish, etc.—is used to make it. Consequently, one could say that to mention expressions is a special case of using them. However, logical analysis requires considering **L** and **ML** as different linguistic systems. The difference has several aspects. One should be especially stressed at this stage, namely that mentioning expressions of **L** in **ML** consists in using their names. In other words, if we refer in **ML** to an expression *E* belonging to **L**, we use (in **ML**) the form '*E*'. Thus, the expression '*E*' functions as the metalinguistic name (it belongs to **ML**) of the expression *E* which belongs to **L**.

Note the difficulty in formulating such statements in ordinary language. In fact, this paragraph simultaneously speaks about two languages, **L** and **ML**. So, several previous statements belong to **MML** (the metalanguage of **ML**). In fact, the expression ''*E*'' (single quotes are twice applied) should be used as the name of '*E*' if we operate in **MML**. In order to avoid artificialities, we adopt the convention (see Preface) that the context 'the expression *E*' replaces (abbreviates) the phrase 'the expression '*E*'' ('it ends with double occurrence of'). Anyway, in-quoting (applying quotes) functions as the device resulting in forming names of expressions. Thus, single-quotations marks can be regarded as a name forming operator (functor) according to the theory of syntactic categories, and its argument is a sentence or other expression. Using of single quotes is, of course, not obligatory. Other conventions (mentioned in previous chapters), like "*E*", <*E*> or $\ulcorner E \urcorner$ occur in logical writings (I will occasionally use the last); we can also use italics (this convention is adopted in his book for writing non-English words as in the sentence '*Schnee*' is a German word referring to snow or other devices, for example capital letters).

(**DG3**) The names 'object-language' and 'metalanguage' are standard now. They became popular owing to the writings of Tarski published in the 1930s, but both probably were invented by Leśniewski (there are no written sources; we have only the oral tradition of the Polish school of logic).►

(**DG4**) In-quoting is not the only way to form names (in-naming) of expressions in **ML**. In fact, Tarski was not very happy with in-quoting (see Tarski 1933, pp. 159–160). He pointed out that the expression '*A*' functions as the individual name for the letter *A*. Consequently, the formula $\ulcorner$'*A*' is true if and only if *A*$\urcorner$, appears as a piece of a grammatical nonsense. Moreover, this version of **T**-scheme is not suitable for generalizations (using the universal quantifier), because nothing can be substituted for *A* in '*A*'—the letter *A* in the latter expression is bounded by the in-quoting operator (see below). The earlier proposed reading (see **DG7III**) the sign '*A*' is a name of a sentence represented by the metavariable *A*, because it does not commit us to regarding the expression '*A*' as a individual name of the letter *A*, but it does not explains the problem of generalization.►

Tarski proposed to use so-called structural-descriptive names, for instance the expression consisting from the letters Es, En, O, Double (in this order) is a name of the word *snow* (italic is an auxiliary device; see above). However (it is my remark), this convention leads to some doubts. Firstly, the role of commas is unclear. If they are dropped, we obtain Es En O Double, but the meaning of this inscription is far from being transparent. Furthermore, the word *snow* like a device similar to in-quoting requires a convention that italicized expressions refer to normally written ones. Tarski usually captured **T**-scheme by the formula

(*) X is true if and only if p,

where X represents a structural-descriptive name (in **ML**) of the sentence p which belongs to a given language **L** (Tarski did not use metavariables). This is OK assuming that the concept of structural-descriptive name has a satisfactory definition. If $E \in \mathbf{L}$, the best way of forming its name in **ML** proceeds via arithmetization, which consists of defining the arithmetic code (see more in Chap. 8, Sect. 8.3) of E, but, unfortunately, this technique has no use outside languages in which arithmetic is representable. Consequently, if **L** does not admit arithmetizations, one must employ other devices. I see no reasons against in-quoting provided that suitable conventions are established in advance. First of all, the letter A in the right part of (e) belongs to **ML** as a metavariable. This is only a minor circumstance, although it clearly indicates that the entire **T**-scheme functions in **ML**. A more important constraint says that the phrase 'A' is the name of a sentence to which A refers (represents). As I earlier remarked, it blocks the symbol 'A' from being the name of the letter A. Another possibility is to use the symbol nE (or any other conventional one) with the proviso that this symbol refers to an outcome of in-naming the expression E (hence, the meaning of this label can be displayed by 'the name of the expression E'). However, I will use the context 'A' (in the explained sense) in further considerations. Note, however, that in-quoting is necessary in mentioning expressions consisting of ordinary words, for example 'snow is white' (the phrase 'the sentence snow is white' is ungrammatical), but can dropped in symbolic expressions, for instance 'the sentence A', which is correct.

I will adopt the following version of **T**-scheme (it is still a preliminary version)

(**TS**) 'A' is true if and only if A.

It allows to the problem of (mentioned earlier) of generalizations of **T**-scheme. Is it possible to use the formula

(**) For any A, 'A' is true if and only if A,

in a coherent way? To repeat, this quantification is not admissible, because the letter A in the context 'A' is not a free variable and is not subject to quantification (also existential). Similarly, the letter X in (*) in (**DG4**) cannot occur under quantification. Consequently, the formula (**TS**), although it is intended as universally valid principle of **STT**, cannot be expressed by (**). Yet every concretization of (**TS**) is generated from it by substituting a name of concrete sentence occurring at the place indicated by the metavariable A.

(**DG5**) Some authors (see Soames 1999, pp. 42–46) propose to work with the substitutional quantification—indeed the objectual one—in order to justify assertions like (**). Roughly speaking, substitutional quantification consists in quantify over variables representing names, not over variables representing objects (as in **FOL**). I will not discuss this strategy.

What about (e) in deriving **LP**? If in (2) we replace the letter l by the sentence 'l is false', we obtain three equalities:

(6) l ='l is false' is false,
(7) $l = l$ is true.
(8) l is false = l is true.

Now, (8) is contradictory (once again, if two sentences are equal, they are also equivalent. Yet the status of the letter l is unclear as before. Once it refers to a sentence about a sentence, but it also functions as a name of the sentence 'l is false'. However, applying (**TS**) to (2) does not resolve **LP**. We have

(9) 'l' is true and only if l.

However, this instance of **T**-scheme, immediately entails

(10) 'l' is true if and only if l is false,

but, because the letter l should be in-quoted in the right part of (10), we obtain

(11) 'l' is true if and only if 'l' is false,

which is contradictory, assuming classical logic and the principle 'A is not true if and only if $\neg A$ is false'.

The above argument shows that the distinction of use and mention does not suffice to eliminate **LP**. As Tarski observed, we must resign from self-referentiality, because, together with **T**-scheme, it entails **LP**. Speaking more precisely, the sentence (L) attributes to itself the semantic property 'is false'. Thus, (L) appears as an object-language utterance, that is, it belongs to **L**. However, the predicate 'is false' belongs to **ML** and thereby can be predicated in assertions pertaining to sentences of **L**. The same applies to the predicate 'is true', which is also semantic; in general, if P is a semantic predicate, not-P shares exactly the same fate. Consequently, **T**-scheme (in its proper, that is, metalinguistic formulation) constitutes a part of **ML**. In the logical syntax forced by the Leśniewski–Tarski diagnosis of semantic paradoxes and resolution of them, the sentences (1) and (2), as well as (6)–(8), cannot be formulated, because they do not satisfy explicitly stated regulative principles of the accepted (in the case) logical grammar. According to Tarski, all semantic paradoxes, including the Grelling antinomy, can be resolved in an analogical manner. To conclude, if $E \in$ **L**, its semantic properties must be investigated and established in sentences (statements) belonging to **ML**. Otherwise, paradoxes are inevitable, unless other remedies (changing logic, modifying **T**-scheme, etc.) are applied.

Yet several serious problems still remain to be discussed. Let me mention two. First of all, some self-referential sentences are not paradoxical—for instance, 'This sentence is written in English'. It is trivially true. On the other hand, the statement 'This sentence consists of seven words' is false, because it consists of six words. Denials of both sentences—'This sentence is not written in English' and 'This sentence does not consist of seven words'—also have definite logical values—the latter is true, but the former is false. The sentence 'This sentence is unprovable' provides an example of a non-trivial self-referring sentence that is not contradictory (see Chap. 8, Sect. 8.4 for further remarks). Hence, the question arises whether there is a general criterion that would separate "dangerous" self-referentialities from sound ones (not producing paradoxes). The issue is not obvious, because the predicate 'is true' behaves, so to speak, correctly, that is, does not lead to a paradox. In particular, the Truth-Teller sentence 'This sentence is true' does not lead to any logical difficulty. On this level of analysis, no answer is possible (see Chap. 8, Sect. 8.6 for further remarks). The second problem revolves around the status of (L). Is it ungrammatical, nonsensical and/or incorrect in yet other sense? Łukasiewicz claimed that it should be banished from logic for its inconsistency. More precisely, he argued that (L) cannot be the value of a propositional variable. However, this view can be objected. We understand (L) as well as other self-referential sentences. Moreover, if the sentence 'This sentence is true' is intelligible, why would its negation have the opposite status. Pointing out that the sentence (L) generates inconsistency does not suffice, because contradictions are fairly intelligible; otherwise, we could not identify them as just inconsistencies. Finally, since the sentence (L) begins a definite sequence of deductive steps performed by standard logical inferential rules, this sentence must be somehow understood according to its meaning. Consequently, we should distinguish (L) as puzzling from obvious syntactic errors, like 'London is if'. However, some problems still remain. What about the sentence 'The number 4 walks slowly'. It is a sentence with so-called category mistake (numbers are not objects which can walk). Is it true or nonsensical? Husserl proposed to distinguish *Unsinn* (nonsense; for instance, 'The number 4 walks slowly') and *Widersinn* (counter-sense; for instance, 'London is if'). Łukasiewicz considered self-referentialities as instances of the *Widersinn*), but Tarski (at least I guess so)—as closer to the *Unsinn.*

A general theory of self-referential sentences is a very controversial issue (see Smorynski 1985 and collections Bartlett, Suder 1987, Bartlett 1992, Bolander, Hendricks, Pedersen 2006 for survey of some proposals in this respect), and I am quite sceptical about a general theory of linguistic self-referring phenomena. It seems to me that the minimal starting condition for a reasonable account of self-referentials from the logical perspective should point out that logic forces some more or less conventional rules in order to resolve difficulties stemming from self-referring of expressions. As far as the issue concerns, the sentence (L) is ill-constructed from the point of view of logical grammar, although its correctness in the light of ordinary grammatical principles is beyond question. Since ordinary language, according to Tarski (see Sect. 7.2, and below) freely combines the use and mention of expressions, it does not distinguish **L** and **ML** sufficiently sharply. Various semantic paradoxes appear as a result of this situation.

(**DG6**) It is true that logical grammar (or syntax, if one prefers) corrects the ordinary grammar of languages. Some authors (in particular, so-called ordinary language philosophers; I skip personal and bibliographical references) argue that colloquial language is perfectly adequate for acts of human communication and even for most scientific tasks, and does not need to be improved (or regimented) by devices of formal logic. However, logical theory does not conform to all characteristic features of ordinary language and cannot do that. The contrary view is akin to advocating that the equation $c + v = c$, absolutely fundamental to special relativity, should be rejected for its plain inconsistency with "ordinary" physics or even Newtonian (classical) physics. Clearly, there is a big metaphilosophical problem concerning the application of logic to philosophy. The considerations in the present book are based on the view that logical means are useful in philosophical analysis, but I am aware that some philosophers contest this position. And, in order to return to an earlier example, the sentence *Das Nichts nichtet*, is fully legitimate.►

(**DG7**) One can formulate several Liar-like paradoxes. One of them (see Woleński 2016), concerns the concept of analyticity. Assume its falls under the scheme

(*) $A \in \mathbf{AN}$ if and only $\mathbf{C}(A)$,

where the letter **C** refers to a condition to be satisfied by any sentence qualified as analytic. For instance, **C** can state 'is true in virtue of meanings', 'is a tautology', 'is true in all possible worlds', 'is true on the basis of the rules of a given language', 'is derivable solely on the basis of logic and definitions', 'has a contradictory negation', etc. In order to make things more explicit, I assume:

(i) A is analytic and true if and only if $\neg A$ is analytic and false;
(ii) **T**-scheme.

Intuitively, (i) asserts the following presumably non-controversial facts: (a) the concept of analyticity is a semantic one (analytic sentences are either true or false); (b) analyticity is closed under negation, that is, denials of analytic sentences are also analytic.

Consider the sentence

(S) (S) is not analytic.

The sentence (S) asserts (about itself) that it is not analytic. Thus, (S) uses the predicate '(not) analytic' self-referentially. Suppose that (S) is true. By (ii), this assumption entails (S). Thus, (S) is not analytic. On the other hand, ¬(S) is false and, thereby, analytic. By (i), (S) is analytic as well. To sum up, if (S) is true, it is analytic and not analytic. Contradiction! Let us suppose now that (S) is false. Thus, it is analytic. By (i), ¬(S) is true as well as an analytic. This implies (by the theorem of **PC**: $(A \wedge B) \Leftrightarrow (A \Leftrightarrow B)$) that $\neg(S)$ is analytic if and only if it is false. Consequently, (S) is analytic if and only if it is true. However, the last assertion entails that (S) is not analytic if and only if it is false. This conclusion is at odds with (i). Contradiction! What is very interesting in this reasoning is the role played by **T**-scheme. It confirms the diagnosis that this scheme plays a very significant role in producing semantic paradoxes.►

(**DG8**) Assume that we use **D**-logic (dual logic). One can prove (see Woleński 1995) that the Truth-Teller sentence leads to a paradox, whereas the Liar sentence does not. Moreover, this fact additionally suggests that one should be very careful with diagnoses of how and why self-referentiality is dangerous.►

(**DG9**) Since I am dealing with **STT**, I do not need to analyse other attempts to re solve **LP**. There are several proposals (see Martin 1984, Barwise, Entchemendy 1987, McGee 1991, Yaqūb 1993, Simmons 1993, Maudlin 2004, Priest 2006, Beall, Armour-Garb 2009 for general as well as concrete suggestions; Visser 2011 offers a valuable survey), including (I do not list all proposals; Chap. 4, Sect. 4.8) truth-value gaps (paradoxical sentences have no truth-values), many-valued logic (paradoxical sentences have other truth-values than truth or falsehood), paraconsistency (paradoxical sentences are true and false at the same time), the revision theory of truth (truth is a circular concept) or the partial definability of the truth-predicate (some instances of **T**-scheme are excluded). Some proposals are ineffective without further ado owing to the Strengthened Liar (this Liar-like Paradox is produced by the sentence 'This sentence is not true'). Moreover, all suggested solutions cost something. For instance, changing logic or modifying **T**-scheme prompts certain reservations. Metaphorically speaking, just as there are constantly tradeoffs in life, so too in constructions of logic. So, opting for any particular solution of the Liar Paradox, must involve weighing both advantages and disadvantages. Take for example, paraconsistent solution. The Liar sentence is considered as an example of a *dialetheia* (a sentence which is simultaneously true and false). Well, but there is a question of distinguishing between *dialetheias* and other inconsistencies. The second objection is this. Defenders of the solution via paraconsistency argue that changing logic blocks logical overfullnes (or explosion), that is, possibility of inferring every sentence from a contradiction. However, it is rather difficult to find an example of using the theorem of **PC** $A \wedge \neg A \Rightarrow B$ in order to deduce an arbitrary sentence from it. I think that Tarski was right arguing that the danger of inconsistency consists in that accepting A and $\neg A$ forces asserting some false statements. This is one reason that, according to my view, Tarski's recipe is quite simple and, moreover, has a very intuitive formal counterpart in metamathematics (see Chap. 8, Sect. 8.6). It is possible that Tarski was influenced by Łukasiewicz at this point. According to the latter (see Łukasiewicz 1910), although the principle of contradiction is not a fundamental logical rule, we should accept it in order to have a weapon against accepting false statements.►

7.5 Language, Metalanguage, Truth and Meaning

The **L/ML** distinction is interesting in itself, independently of its role for **LP**. Since speaking of **ML** involves **MML**, and so on, we have the hierarchy **HL** = $\mathbf{L_o}$ (the object-language), $\mathbf{L_1}$ (= $\mathbf{ML_o}$), $\mathbf{L_2}$ (= $\mathbf{ML_1}$), $\mathbf{L_3}$ (= $\mathbf{ML_2}$), …, $\mathbf{L_n}$ (= $\mathbf{ML_{n-1}}$), …. This hierarchy is infinite—if we reach the stage k ($k \geq 0$), passing to $k + 1$ is straightforward. Now, defining semantic properties of $\mathbf{L_k}$ requires using $\mathbf{L_{k+1}}$. Two questions immediately arise. Firstly, what about relations between $\mathbf{L_k}$ and $\mathbf{L_{k+1}}$ (= $\mathbf{ML_k}$)? According to Tarski (Tarski 1944, p. 675):

> The solution [of the problem truth-definition – J. W.] […] depends upon some formal relations between the object-language and tits meta-language; or, more specifically, upon the fact whether the meta-language in its logical part is "*essentially richer*" than the object-language or not. It is not easy to give a general and precise definition of this notion of "essential richness." If we restrict ourselves to languages based on the logical theory of types, the condition for the meta-language to be "essentially richer" than the object-language is that it contains variables of a higher logical type than those of the object-language.

For simplicity, some of my further remarks omit the indices of languages and refer just to **L** and **ML**. According to Tarski, **ML** must contain **L** (Tarski 1944, p. 674):

> […] every sentence which occurs in the object-language must also occur in the meta-language; in other words, the meta-language must contain the object-language as a part. This is at any rate necessary for the proof of the adequacy of the definition […].
>
> (The requirement in question can be somewhat modified, for it suffices to assume that the object-language can be translated into the meta-language. […]. In all that follows we shall ignore the possibility of this modification).

As we see from the bracketed passage, Tarski preferred the latter view (but earlier, that is, in Tarski 1933, he used the former; see also below). This can be explained by the comments (I use in-quoting) on the formula '*X* is true if and only if *p*' (see (**DG4**)). The phrases '*X*' and '*X* is true' belong to **ML**. On the other hand, $p \in$ **L**. Finally, (*) $\in$ **ML**. Since the predicate 'is true' appears in **ML**, not in **L**, including the latter in the former does not lead to antinomies. Yet one could claim that **T**-scheme should be syntactically uniform in the sense of being formulated entirely in **ML**. How to fulfil this requisite? Turning to the (e) (the refined **T**-scheme), we must, as Tarski suggested, modify its right side. Let the symbol A* means 'a translation of an expression represented by the letter *A i*nto **ML**'. Using this symbolism, we obtain the formula

(**TS***) '*A*' is true if and only if A*,

which is the final version of **T**-scheme adopted in this book. Further details depend of how translations of **L** into **ML** are prescribed. To anticipate Chap. 8, Sect. 8.2, I give an example:

(12) The sentence 'snow is white' is true if and only if snow $\in$ **White**,

where the word '**White**' refers to the set of white objects. This translation indicates that set theory (more precisely, the Zermelo–Fraenkel system but without taking into account its various theoretical problems) serves as the translation manual for **STT**. To conclude the comments on the relation of **L** and **ML** let me add that logical constants (connectives and quantifiers) are invariants over the whole of **HL** (this remark is a trivial supplement to Tarski's view mentioned in (**DG16V**), or can be considered as an application of a principle of the theory of logical types).

(**DG10**) Tarski in his preference for including **L** in **ML** as opposed to translating the former into the latter, was probably guided by his nominalistic sympathies. Translating is an operation that invokes directly the concept of meaning, which Tarski tried to avoid. For instance, he urged (see Tarski 1936b) that definitions of semantic concepts, including that of truth, should be relativized to a language, not to meanings; these remarks were addressed to Kokoszyńska who proposed (see Kokoszyńska 1936) relativizing to meanings of expressions in a given language. He regarded the concept of language as simpler than the concept of meaning. Kokoszyńska replied that Tarski's answer begs the question, because we have to do with interpreted languages.►

The next issue concerns the possibility of the universal **ML**, say, **UML**. It is ruled out if antinomies are to be avoided. More precisely, $\mathbf{L_0}$ is the object-language. Its semantic theory, say, $\mathbf{ST}(\mathbf{L_0}) = \mathbf{ST_1}$ finds its wording in $\mathbf{ML_0}$ (= $\mathbf{L_1}$). Unbounded repetition of this step yields an infinite sequence (hierarchy) of semantic theories $\mathbf{ST} = \mathbf{ST_1}, \mathbf{ST_2}, \mathbf{ST_3}, \ldots$ (note that $\mathbf{L_0}$ is not enough for defining truth for it and can never be such. Assume that the language of the universal semantic theory plays the role of such **UML** in which we can say something about all theories including statements about **ST**(**UML**). Clearly, this language is semantically closed and produces inconsistency.

(**DG11**) Following Putnam 1975a, pp. 72–73 (see also Wang 1986, p. 143), I prefer to speak rather about theories, not languages. Every language is contradictory if it includes the symbol $\neg$ and the rule that if A is a formula, so is $\neg A$. Consequently, it is actually important to investigate whether theories—that is, sets of sentences—are consistent or not. In the case of **LP**, this investigation in fact involves a small portion of ordinary language. On the other hand, every language is a set of sentences and can be investigated, assuming (**BI**) as consisting, of two subsets, truths and falsehoods. This circumstance justifies, at least to some extent, Tarski's approach.►

Note, however, that **UST** is possible via changing its logic from the classical system to some other (see (**DG9**)). The price of such a manoeuvre is to abandon the idea that the meaning of logical constants is the same throughout the entire **HL** and **ST**. Yet one can ask, 'What about the statements about **HL** and **ST** in this book?' In particular, do they not look like general (universal) assertions applied to every item of $\mathbf{L_n}$ belonging to both hierarchies. That is true, but they do not violate the constraints that need to be imposed on a language in order make it to be free of antinomies, because they are not belonging to **UST**. Of course, the last statement is an empirical generalization, but—employing the legal slogan *in dubio pro reo*—an assertion is innocent to be antinomy-generating as long as it can be proved that its use produces a contradiction.

(**DG12**) The situation of **HL** (and **UST**, but I will omit noting this in subsequent remarks) is very similar to that of the hierarchy of logical types (see Chap. 6, Sect. 6.3) or of a set-theory constructed as a cumulative hierarchy of sets. In all these cases, we constantly encounter similar or analogical questions. Is the universal metalanguage possible? Is possible the logical type of all types? Is possible the set of all sets? And in all of these situations, the answer is analogous, namely 'Yes', but provided that **UML**, the logical type of all types, or the set of all sets behave as abnormal objects: a language with different properties than standard languages, a logical type defined in a peculiar way, for instance, by a restriction of logic in reasoning about this peculiarity (it turns to restrictions on **UML**-logic) or as class which is not a set (it forces that classes to have no over-classes). Some philosophers probably would say that such very peculiar objects are transcendental in their relation to customary languages, ordinary logical types or typical sets. I think that it is rather nice that such different constructions have to meet similar challenges in order to be subjected to consistent reasoning about them, or provoke similar philosophical discussions concerning their ontological status. It is an open and interesting question whether various proposals concerning set-theory, for instance, the von Neumann–Gödel–Bernays set theory (with classes as genuine objects) or one that employs the category theory as the foundation of mathematics, could essentially change conclusions pertaining to **HL**. As far as I know, new special restrictions must be introduced sooner or later—for example, that not all set-theoretical operations are performable on classes, or that so-called small categories are admitted. Finally, it is interesting (see Church 1976) that Tarski's method of coping with semantic antinomies is more general than that of Russell, because does not require using tools similar to that introduced by the ramified theory of types.►

Yet we should address the problem of the possibility of a logical analysis of ordinary language (see also Fenstad 2004). Tarski's initial view on this issue is captured by the following quotation (Tarski 1933, pp. 164–165):

> A characteristic feature of colloquial language (in contrast to various scientific languages) is its universality. It would not be in harmony with the spirit of this language if in some other language a word occurred which could not be translated into it; it could be claimed that 'if we speak meaningfully about anything at all, we can also speak about it in colloquial language'. If we are to maintain this universality of everyday language in connection with semantical investigations, we must, to be consistent, admit into the language, in addition to its sentences and other expressions, also the names of these sentences and expressions, and sentences containing these names, as well as such semantic expressions as 'true sentence', 'name', 'denote', etc. But it is presumably just this universality of everyday language which is the primary source of all semantical antinomies […].
>
> If these observations are correct, then the *very possibility of a consistent use of the expression* '*true sentence*' *which is in harmony with the laws of logic and the spirit of everyday language seems to be very questionable, and consequently the same doubt attaches to the possibility of constructing a correct definition of this expression.*

This negative attitude was weakened in Tarski 1944, pp. 670–671. In particular, Tarski (see p. 670) introduced the concept of languages with specified structure:

> There are certain general conditions under which the structure of a language is regarded as *exactly specified*. […] to specify the structure of a language, we must characterize unambiguously the class of those words and expressions which are to be considered *meaningful*. […] we must set up the criteria for distinguishing […] "*sentences*". Finally, we must formulate the conditions under which a sentence of the language can be *asserted*. […] we must indicate all *axioms* […] and […] so called *rules of inference* […] by means of which we can deduce new asserted sentences from other sentences which have been previously asserted.

Recent studies (see proposals in Kashtan 2017) show that application of formal tools to analysis of the concept of truth in ordinary language can be considerably increased. In general, although formal semantics of natural languages is still problematic as a general theory, its partial realizations are more and more successful.

A language is formalized if its description appeals exclusively to the form of its expressions (Tarski 1933, p. 165/166):

> These can be roughly characterized as artificially constructed languages in which the sense of every expression is unambiguously determined by its form.

However, not every specified language is of this kind (Tarski 1944, p. 671):

> […] we can imagine the construction of languages which have an exactly specified structure without being formalized. In such a language the assertibility of sentences, for instance, may depend not always on their form, but sometimes on other, non-linguistic factors. […].
>
> *The problem of the definition of truth obtains a precise meaning and can be solved in a rigorous way only for those languages whose structure has been exactly specified.* For other languages—thus, for all natural, "spoken" languages—the meaning of the problem is more or less vague, and its solution can have only an approximate character. Roughly speaking, the approximation consists in replacing g a natural language (or a portion of it in which we are interested) by one whose structure is exactly specified, and which diverges from the given language "as little as possible".

The new element consists in admitting a portion of ordinary language as suitable for a precise semantic analysis, assuming that this selected linguistic segment is specified. This fact does not exclude that such "prepared" language is similar to its version before specification. Yet ordinary language in its full totality is universal (in the sense mentioned above) and thereby inconsistent, but, on the other hand (see Davidson 1967), nobody uses ordinary language in its entirety. If so, we can always extract a portion of everyday parlance and submit it to semantic analysis in Tarski's sense.

(**DG13**) In fact, Tarski anticipated his view on specified languages and their role for ordinary language in Tarski 1933, p. 165, note, when he was speaking on partial applicability of results obtained for formalized languages for formulation of fragmentary truth-definition "which embraces a wider or narrower category of sentences". Yet introducing a concept of specified language seems very important.

Tarski had substantive reasons (see Woleński 2017a) for changing the Polish title of Tarski 1933, *Pojęcie prawdy w językach nauk dedukcyjnych* (The Concept of Truth in Languages of Deductive Sciences) into *Der Wahrheitsbegriff in den formalisierten Sprachen* (The Concept of Truth in Formalized Languages). Due to his various deliberations, he came to the conclusion that the original Polish title did not reflect his views about language (see Patterson 2012 for a comprehensive analysis of Tarski's philosophy of language). It seems he decided that the label 'languages of deductive sciences' is too vague and should be replaced by 'formalized languages'.►

Yet even formalized languages have a connection with what is not formal. Tarski strongly underlined this circumstance (Tarski 1933, p. 166/167):

> It remains perhaps to add that 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 material sense is attached. For such sciences the problem here discussed [the problem of truth—J. W.] 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. The expressions which we call sentences still remain sentences after the signs which occur in them have been translated into colloquial language. The sentences which are distinguished as axioms 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.

A more general view appears in the following remarks (Tarski, Givant 1987, p. 1, 18/19, 22):

> Axiomatic systems of set theory developed in the formal language **Ł** of the (first-order) predicate logic with identity and in some other languages with different formal structures are the central topic of the present monograph. […]. The discussion is conducted throughout the book within an appropriate *metasystem.* In the *metalanguage*, i.e., the language of the metasystem, we have at our disposal various logical, set-theoretical, and metalogical symbols and notions. […]. The metasystem and its language are not assumed to be formalized. The set-theoretical notions occurring in the metasystem are sometimes employed in a way which is usually described by the phrase "in the sense of naive set-theory". […]. Among metalogical notions of the metasystem we find, in particular, symbolic designations of all expressions occurring in formal languages to which the discussion refers. No symbols, i.e., expressions appearing in our metalogical discussion, should be interpreted as belonging to formal languages themselves. […]. In this work we use the terms "formalism" and "formal language" […] interchangeably. In other contexts it may be useful to differentiate between the meanings of these two terms. Formal languages would then be constructed as structures with a different list of fundamental components; the list would include some notions referring to the intrinsic structure of sentences such, as the vocabulary of a language. […].
>
> There is another notion of a general character, closely related to the notion of formalism, that will frequently be used in this work, namely the notion of a system. […]. Actually, for purposes of this work, we restrict ourselves to those systems which are developed in a formalism or, what amounts to the same thing, to systems obtained by relativizing a formalism to a certain set of sentences. Thus, we shall speak of the system of Zermelo set theory as a system […] obtained by relativizing **Ł** to a well-known set of Zermelo's axioms. Similarly, Peano arithmetic can be referred to as a system developed in a first-order

> formalism (with appropriate nonlogical constants), or else as a system obtained by relativizing this formalism to the set of Peano's axioms. [...]. We shall only use the term "system" in application to interpreted formalisms.

Indeed, relativizing formalisms to concrete sets of axioms just causes deductive systems to be interpreted.

For Tarski, doing formal semantics for **L** requires that this language be formalized and interpreted. This account certainly leaves important questions open. For example, one can ask what it means that the sense of every expression occurring in a formalized language is unambiguously determined by its form, even if we take into account that this explanation should be taken *cum grano salis* or approximate. Leaving aside a deeper discussion—which would by a conceptual necessity have to go the difficult question concerning the concept of meaning—it seems that the most important lesson afforded by the just quoted passages from Tarski (at least a discussion that revolves around **STT**), is that the adjectives 'formalized' and 'interpreted' can perfectly well coexist as attributes of languages. That lesson appears as quite obvious when one bears in mind that 'formalized' refers to the outcome of the process of formalization. I regard Tarski's remarks on formalization and interpretation as of the utmost importance for the proper understanding of **STT** (see Chap. 8, Sect. 8.2). Once again (see Chap. 6, Sect. 6.6), we encounter here a very good example of interplay between what is formal and what is informal.

The last problem to be discussed in this section concerns the relativization of a truth-definition to a given language **L**. In other words, we define not the predicate 'is true' but the more complex expression 'is true in **L**'. This restriction is very frequently explained by the following kind of example:

(13) The sentence *Schnee ist weiss* is true if and only if snow is white.

In (13), German serves as **L**, but English as **ML**. Proceeding further, we also have here an example of translation, because the English sentence 'snow is white' translates the German sentence *Schnee ist weiss* from **L** to **ML**. Yet this illustration is misleading, or at least may lead to confusions. In fact, it only shows how the process of translation operates, if this option works as the chosen way of correlating both **L** and **ML**. There is no difficulty in saying that the sentence *Schnee ist weiss* occurring in German as the object-language got integrated into English function as the metalanguage. Nothing new shows up when we use the same ethnic language for **L** and **ML**. Consider

(14) The sentence 'snow is white' is true if and only if snow is white.

This entire formula belongs to **ML**, but the right part of (14) can be regarded as an English sentence of **L** included in **ML**. This is a natural approach, but to say that the sentence occurring as the right part of (14) is translated into **ML** requires rather the pronounced artificiality of its self-translation into English as the relevant metalanguage. Incidentally, to say that the German sentence *Schnee ist weiss* is included in the English as **ML** does not sound all too natural. We have here a good lesson that grammatical rules of ethnic languages do not assimilate some constraints imposed on languages that are formalized or specified (in Tarski's sense).

What is going on with 'is true in **L**'? The answer is very simple. The letter **L** functions as **a** parameter indicating that truth is defined with respect to a language of a definite level in the **HL**. Strictly speaking, we should say 'is true in $\mathbf{L_n}$' instead 'is true in **L'**, but because the issue usually concerns $\mathbf{L_o}$, dropping the index does not result in any misunderstandings. The parameter **L** immediately, although indirectly, invokes **ML** as the homeland for defining the concept of truth. Of course, **L** and **ML** must satisfy further conditions related to their syntactic features (**L** must be formalized or at least specified, but also interpreted, and **ML**—based on a clear manual of translation, etc.). Take for example $\mathbf{L_0}$, $\mathbf{ML_0}$ (= $\mathbf{L_1}$) and $\mathbf{ML_1}$ (= $\mathbf{L_2}$). Assume that we define (a) 'is true in $\mathbf{L_0}$', and (b) 'is true in $\mathbf{L_1}$'. Are both predicates the same or not? This problem is quite controversial, for it looks like **STT** converts the homogeneous concept of truth into a family of different notions associated with particular levels of **HL**. I will return to this question in Chap. 4, Sect. 4.2.

7.6 Heuristics and the Conditions of Adequacy with Respect to SDT

As I noted earlier Tarski considered the concept of satisfaction (more precisely, the satisfaction relation) as basic in semantics. As far as the issue of truth-definition is concerned, he decided to define truth as a special case of satisfaction. This approach can be reconstructed as follows (I am not suggesting that Tarski's actual heuristics was exactly that). Open formulas are defined in $\mathbf{L_{FOL}}$ (the language of **FOL**) as containing free variables (see Chap. 5, Sect. 5.2.3). By contrast, closed formulas have no free variables—for instance, $P(a)$ or $\exists xPx$. Open formulas are satisfied or not, depending how the free variables are interpreted by a given valuation function, but sentences are true or false. The following example plays an illustrative role. Let **U** be a universe, say, that of natural numbers. Since our considerations at this point are elementary and intuitive, we do not need to consider the entire model associated with **U** (I use in this fragment concepts defined in Chap. 5, Sect. 5.2.4). Consider the formula

(15) x is a prime number in **U**,

which is satisfied, if, for instance, $\mathbf{v}(x) = 3$, but not satisfied, if, say, $\mathbf{v}(x) = 4$. Truth and satisfaction are connected in the case of (15) and similar formulas in the sense that the valuation function ascribes to (free variables) objects which belong to the scope of a given predicate, for instance 'is a prime number'. If an object **a**, denoted by the constant a, satisfies the formula $P(x)$, the sentence $P(a)$ is true.

Roughly speaking, satisfaction converts formulas into true sentences, but non-satisfaction into false ones. Consider now the formula

(16) $\exists x$(x is a prime number).

Its truth depends on what is substituted for x in the expression $P(x)$ in the sense that the truth of (16) is inferable (or not) from 'a is a prime number'. Consequently, if $\mathbf{v}(x) = 3$, (16) becomes true. However, if $\mathbf{v}(x) = 4$, (16) is converted into a false sentence. This substitution does not provide a basis for deriving (16) from '4 is a prime number'. In general, the sentence $\exists xP(x)$ is false in the case that no object satisfies the formula $P(x)$. For example, the sentence

(17) $\exists x$(x is the greatest natural number),

is demonstrably false, because in **U** no object is the greatest natural number. Consider now the sentence

(18) $\forall x$(x is a natural number).

The formula 'x is a natural number' is satisfied by any element taken from **U**, because, by definition, it consists exclusively of natural numbers and nothing else. Consequently, the sentence (18) is demonstrably true. In fact, **FOL** establishes that if $P(x)$ holds (is satisfied) by any object, then its universal closure $\forall xP(x)$ is true, and that if Px is satisfied by at least one object, $\exists xP(x)$ is true. Thus, (at least some) links between satisfaction and truth are generated by pure logic.

The above considerations do not provide a definition of truth. Consider now two collections of ideas:

(19)

(General case):	open formulas,
	satisfaction by some objects from **U**;
	non-satisfaction by some objects from **U**;
(Special case):	closed formulas (sentences), satisfaction by?;
	non-satisfaction by?

Returning to (17) and (18), each assertion provides a case in which logical values of sentences depend, so to speak, on the behaviour of the entire **U**, that is, content of the assertions in question is represented by their objects and the most general properties as related to the totality of objects in **U**. More specifically, the formula $\forall xP(x)$ is true, if $P(x)$ is satisfied by any $\mathbf{a} \in \mathbf{U}$, but $\exists xP(x)$ is false if $P(x)$ is not satisfied by any $\mathbf{a} \in \mathbf{U}$. On the other hand, the truth of the formulas $\mathrm{P}(a)$ and $\exists xP(x)$ as well as the falsity of $\forall xP(x)$ cannot be explained by satisfaction in a simple way (of course, this conclusion holds just as well for more complex formulas, for instance, $\forall x\exists yP(xy)$). These remarks show that the relation between satisfaction and truth has many dimensions (see also Chap. 8, Sect. 8.2 and Chap. 4, Sect. 4.2).

The last paragraph suggests that we need a generalization for obtaining the required truth-definition based on the concept of satisfaction. Inspecting the special cases discussed leads to the conclusion that although satisfaction depends on valuation of free variables, truth and falsehood do not. The reason is very simple and even trivial, namely that sentences have no free variables. Consequently, truth and falsehood should (even must) be independent of how the valuation function acts with respect to terms that are free variables. On the other hand, logical values are determined by valuations of constants, function symbols, and predicates as well as by the understanding of quantifiers. For instance, if we say that 1 is a natural number, the truth of this assertion depends essentially on the value of the numeral 1 (strictly speaking, numerals should be distinguished from numbers, but this ambiguity does not matter in informal considerations) and the predicate 'is a natural number'. In the cases of sentences (17) and (18), we have a special situation, because the valuations in question appeal to all objects and, in this sense, are independent of mappings of free variables to concrete objects. In other words, whatever object (from **U**) is correlated with x, (17) remains false, but (18)—true.

The last observation motivates the following formulation of **SDT** (see Tarski 1944, p. 677), assuming that the domain of interpretation **U** is fixed:

(20) (a) 'A' is true if and only if 'A' is satisfied by any object in **U**;
(b) 'A' is false if and only if 'A' is satisfied by no object in **U**.

Using (18) as an example, we have $A = \forall x(x \in \mathrm{N})$, where **N** (= **U**) is the set of natural numbers. Now, (19) can be corrected by dropping question-marks as

(21)

Open formulas:	satisfaction by some objects from **U**, but not others;
sentences:	satisfaction by all objects from **U** (truth);
open formulas:	non-satisfaction by some objects from **U**;
sentences:	satisfaction by no objects from **U** (falsity)

This definition (20) is still provisional and will be modified in Chap. 8, Sect. 8.2.

(**DG14**) Tarski (see Tarski 1936, pp. 405–406) also considered also introducing semantic concepts, not by definition, but axiomatically, but according to him:

> […] when this method [axiomatic], which seems easy and simple, is worked out in detail various objections present themselves. The setting up of an axiom system sufficient for the development of the whole of semantics offers considerable difficulties. […] the choice of axioms always has a rather accidental character, depending on inessential factors (such as e.g. the actual state of our knowledge). Various criteria which we should like to use in this connection prove to be inapplicable. Moreover, the question arises whether the axiomatically constructed semantics is consistent. The problem of consistency arises, of course, whenever the axiomatic method is applied, but here it acquires a special importance, as we see from the sad experiences we have had with the semantical concepts in colloquial language. […] this method […] would arouse certain doubts from a general philosophical point of view. It seems to bring this method into harmony with the postulates of the unity of science and of physicalism (since the concepts of semantics would be neither logical nor physical concepts).
>
> In the second procedure, which has none of above disadvantages, the semantical concepts are defined in terms of the usual concepts of the metalanguage and are thus reduced to purely logical concepts, the concepts of the language being investigated and the specific concepts of the morphology of language. In this way semantics becomes a part of the

> morphology of language if the latter is understood in a sufficiently wide sense. The question arises whether this method is applicable at all. It seems to me that this problem can now be regarded as definitely solved.

Leaving aside at this point the problem of physicalism (see Chap. 9, Sect. 9.4), Tarski's doubts about using the axiomatic method in semantics seem to stem from his belief, related to the general methodology recommended in the Lvov–Warsaw School, that definitions provide the best method for explaining concepts. Of course, Tarski as a mathematician did not condemn the axiomatic method at all, but he was careful not to overestimate its role—at least outside of pure logic and mathematics. His view that "the problem can now be regarded as definitely solved" must at the moment be qualified as premature, because axiomatic approaches to the concept of truth have been popular in recent years (see Halbach 2011, Horsten 2011, Cieśliński 2017).

An important problem appears in the second part of the quotation in this digression. It concerns the status of concepts (expressions) used in **SDT** (more generally, in **STT**). Tarski's claim that these concepts belong to the morphology of **L** and **ML**, can be regarded as intelligible, because it meets an objection that the definition in question is circular (see also Sect. 7.4 and Chap. 9, Sect. 9.2). Yet the question is much more difficult, because **ML** is not formalized. Hence, the labels 'usual', 'specific' or 'sufficiently wide' as applied to **ML** (or even to **L**) have no sufficiently precise sense. Consequently, Tarski's assertion that truth-definition uses notions "reduced to purely logical concepts" is vague.►

The definition of sentences as open formulas without free variables looks at first sight like an artificial mathematical trick, but such constructions frequently occur in mathematical practice as useful simplifications. For example, the straight line can be considered as a special case of a curve, or Euclidean space as a special instance of Riemannian space. Consequently, (20) can be charged with being a result of a purely formal game, completely alien to ordinary and philosophical intuitions. Tarski did not conceal that his explanations pertaining to what truth employ mathematical concepts and techniques perhaps fairly obvious for practising mathematicians, but not convincing as tools of a reasonable philosophical analysis. I have no intention to deny that. However, we can also try to argue that this definition fulfills some intuitive constraints. Since (20a) and (20b) entail that no sentence is true and false at the same time, we obtain the metalogical principle of contradiction. On the other hand, if A is an open formula, it is not the case that either A is satisfied or $\neg A$ is satisfied. The formulas $P(x)$ and $\neg P(x)$ can serve as an example—both can be satisfied, for instance, 'x is a city' and 'x is not a city' This example shows that generally speaking satisfaction of open formulas has some other properties than truth as an attribute of sentences, although, to say once again, both concepts are related in many ways. By definition, every sentence is satisfied by all objects or by no object. Assume that the formula $\forall x P(x)$ is true and, thereby, satisfied by every object. Now, its negation, the formula $\exists x \neg P(x)$, is satisfied by no object. This assertion implies the metalogical principle of the excluded middle. Thus, we reach (**BI**), that is, the principle of bivalence. These facts will be made more precise in Chap. 8, Sect. 8.2. Finally, let me try to come up with a philosophical paraphrase of the statement that if truth and falsehood are independent of valuations of free variables, then having logical values by sentences depends on how things are models, in our examples, in **U**. Perhaps we could say that if truth and falsehood are indeed free of such valuations, then whether sentences have definite logical values of how things are in a relevant model, in our examples on **U**.

(**DG15**) Two additional remarks are in order. Firstly, satisfaction by all objects cannot be regarded as equivalent to being a logical tautology. Satisfaction is always relative to a chosen (fixed) universe. In particular, all conclusions made in this section assume that the stock of predicates—such as 'is a natural number', 'is a prime number', or 'is the greatest natural number'—is established in advance and its elements have a definite meaning that stems from a specific interpretation. If A is a logical tautology of **FOL**, this means that A is true (now in the outlined sense) in all models (see (**Df13c**) in Chap. 5). Secondly, (20)—and this is, a new factor, relativizes truth (and falsehood) not only to **L**, but also to a model **M**. This gives a new meaning to the phrase $\mathbf{M} \models^{\mathfrak{I}} A$ (the sentence A is satisfied (true) in a model **M**, relative to an interpretation $\mathfrak{I}$). This symbolism can be extended to $\mathbf{M} \models^{\mathfrak{I},\mathbf{L}} A$ (the sentence A is satisfied (true) in a model **M**, relative to an interpretation $\mathfrak{I}$ of a language **L**). I will continue these remarks in Chap. 8, Sect. 8.2 and Chap. 9, Sect. 9.2, in particular, about the role of all objects, more precisely, of all sequences of objects in interpreting **SDT**.►

To be satisfactory **SDT** must conform to so-called conditions of adequacy. More specifically, this definition must be (a) formally correct, and (b) materially correct (according to Tarski, every good truth-theory must respects these constraints). Condition (a) was not explicitly stated by Tarski. Clearly, he was thinking that the definition should be consistent, that is, not resulting in antinomies. The requirements involving the interplay of **L** and **ML** function as insurance against semantic inconsistencies. It seems that (a) covers some further traditional constraints—like non-circularity, or overcoming the definition *idem per idem.* Roughly speaking, **SDT** is not circular and does not proceed by *idem per idem*, because it does not assume the concept of truth in **ML** (see also Chap. 9, Sect. 9.2). Condition (b) was stated in such the following way (Tarski 1933, p. 187/188; this fragment is originally written in italic except symbols):

> CONVENTION **T**. A formally correct definition of the symbol 'Tr', formulated in the metalanguage, will be called an adequate definition of truth if it has the following consequences:
>
> (α) all sentences which are obtained from the expression '$x \in Tr$ if and only if p' by substituting for the symbol 'x' a structural-descriptive name of any sentence of the language in question and for the symbol 'p' the expression which forms the translation of this sentence into the metalanguage;
>
> (β) the sentence 'for any x, if $x \in \mathrm{Tr}$, then $x \in \mathrm{S}$' (in other words $Tr \subseteq S$).

Condition (β) of the convention **T** (**CT** for brevity) says that true sentences form a subset of **L** (Tarski used the letter S). Condition (α) claims that any materially correct truth-definition must entail all **T**-equivalences (**T**-sentences), that is, particular instances of **T**-scheme. Note that the first constraint does not claim that the formula (in my notation)

(22) $\forall A$('A' is true if and only if A^*),

is entailed by **SDT**, because, as I have already noted, the letter A cannot be quantified for its occurrence inside the quotes in the formula 'A' is true if and only if A^*. The proper version says that all **T**-equivalences belong to consequences of **SDT**.

(**DG16**) The letter **T** in 'Convention **T**', '**T**-scheme', '**T**-equivalences' or '**T**-sentences', does not allude to Tarski's surname, but just to truth. In the Polish original is *Konwencja P* (Convention *P*; related to the word *prawda* (truth), in German translation—*Konvention W* (from *Wahrheit*).►

The condition of material adequacy, particularly its part (α), shows that **T**-scheme is not a truth-definition. On the other hand, Tarski underlined that every particular **T**-sentence provides a partial definition of truth for a given sentence. One could possibly form the conjunction of all **T**-equivalences as the definition, but this formula would to be infinite in length, and as such pointless; in particular, it does not directly follow from (α). As far as the issue concerns, (**BI**), it cannot be obtained from **T**-scheme. Decomposition (by **PC**) of the formula 'A' is true if and only if A^* gives the conjunction

(23) ('A' is true $\wedge$ A) or $\neg$('A' true) $\wedge$ $\neg A^*$),

but, even, if we simplify it to

(24) 'A' is true or $\neg$('A' true),

this formula is still essentially weaker than (**BI**) (see Chap. 4, Sect. 4.8). In fact, **STT** is richer than **SDT** itself.

This chapter allows answering to some issues connected with the tasks of truth-theories. First of all, **SDT** satisfies all Russell's conditions, if we regard sentences as truth-bearers. So **SDT** defines truth a property of sentences, by their relation to something other than sentences (in other words, as an external relations) and provides a theory of falsity. The predicate 'is true' is a determinator and the same concerns 'is false' (this is a property of epistemological notion of truth). On the other hand, truth is definable by **SDT**, but one can see that **SDT** defines a set of truths in a given language **L**. So **SDT** provides an extensional truth-definition. Using an old intuition that every extension is associated with an intension, we can try to consider what is the latter as related to **SDT**, but the answer is not straightforward. Anyway, **SDT** is a substantive definition of truth; **STT** is such *a fortiori.* Finally, **SDT** leads to definite consequences about the relation between truth and logic. And if we agree that (20) provides a way of understanding how things are, **SDT** conforms to the traditional intuition of the correspondence (better, the classical) theory of truth. Tarski himself linked philosophical aspects of **STT** with **T**-scheme, but this view seems too minimalistic (see Chap. 9, Sect. 9.2 for a more extensive discussion).

(**DG17**) Framing **STT** in frameworks of **FOL** allows an answer to the objection (see Black 1949, Kripke 1975) that the stratification of the concept of truth into several notions related to levels of **HL** is an outcome of Tarski's construction, but, it s artificial from the intuitive point of view, because we use the unitary predicate 'is true' in all practical circumstances (see, however, Kashtan 2017 for a defence of the stratified concept of truth). The ordinary understanding of truth is give by the form '*A* is true', but **STT** generates the hierarchy 'truth in $\mathbf{L_0}$', 'truth in $\mathbf{L_2}$', ... Although this argument overlooks that $\mathbf{Truth}(\mathbf{L_{n-1}}) \subset \mathbf{Truth}(\mathbf{L_n})$, **SDT** must be performed on every level of **HL**. Taking **FOL** as the foundation changes the situation, if we simultaneously accept (**FOT**) and the Hilbert thesis (every theory can formalized in the first-order language), because we have 'true in the first order **L**, which is defined in **ML**. The question whether we need 'is true in **ML'** in order to define 'true in **L**' will be addressed in Chap. 9, Sects. 9.1 and 9.2.►

(**DG18**) I use two labels: **STT** (semantic theory of truth) and **SDT** (semantic definition of truth). In Sect. 7.2, I listed several components of **STT**, and **SDT** is one of them, clearly the most important. Hence, most discussions about **STT** concentrate on **SDT** or even consider them as equivalent. I hope that using of these two denominations does not lead to any misunderstanding.►

(**DG19**) This chapter well illustrates the fundamental role of Tarski's central maxim (see Chap. 6, Sect. 6.6 in **STT**).►

Appendix: Yablo Sequences and Self-Reference

Stephen Yablo (Yablo 1993) produced a version of **LP** in which "self reference is neither necessary nor sufficient. Consider the following sequence of sentence (Yablo uses 'untrue', not 'false'):

(A_1) for all k > *1*, A_k is false;
(A_2) for all k > *2*, A_k is false;
(A_3) for all k > *3*, A_k is false;
....

Consider (I follow Cook 2014, pp. 11–12, but my reasoning is slightly different; see also Kripke (2019a) and Cook 2014, pp. 27–28 for a metamathematical derivation of the Yablo paradox) the sentence A_m (it says that for all $n > m$, A_m is false). Assume that A_m is true. Hence, it is true what A_m says. Thus, $n > m$, A_m is false. In particular, the sentence A_{m+1} is false. However, due to the assumption that A_m true it is impossible, because falsity of A_{m+1} implies that there is sentence A_n ($n > m + 1$) which is true. Consequently, A_{m+1} must be true. Since A_m and A_{m+1} are both true, they are equivalent. Using the formula $A_m \Leftrightarrow$ (A_{m+1} is false), replacing equivalents gives $A_m \Leftrightarrow$ (A_m is false). The last formula immediately leads to **LP**. Assume that A_m is false. Thus, all sentences A_k ($1 < k < m$) are also false. Suppose that A_i and

A_{i+1} are such sentences. Since they are both false, they are equivalent too. Since we have the formula $A_i \Leftrightarrow$ (A_{i+1} is false), replacing equivalents gives the equivalence $A_i \Leftrightarrow$ (A_i is false).

The above argument explicitly shows that self-reference and **T**-scheme are involved into the Yablo Paradox. I do not find a sufficient argument for Yablo's conclusion that **LP** can be formulated without referring to self-reference (a similar view was expressed by Kripke—a personal communication), if not explicitly, then—at least implicitly. His informal reasoning is incomplete; the proof in Cook 2014 although does not appeal to self-reference plays with A_1 is false and A_1 is true without noticing that the paradox arises as a result of asserting own falsity by A_1. In fact, the reduction of A_{i+1} to A_i (for any i) implicitly involves self-reference. Thus, I conclude that the Yablo paradox does not invalidate the Leśniewski-Tarski diagnosis. Further remarks about this issue will be found in Chap. 8, Sect. 8.8. Let me note that I entirely omit the problem of circularity in semantic paradoxes (see Cook 2014 for an extensive discussion of this problem).

References

Abeles, F., Fuller, M. E. (Eds.) (2016). *Modern Logic 1850–1950. East and West*. Basel: Birkhäuser.

Bar-Am, N., Gattei, S. (Eds.) (2017). *Encouraging Opnness. Essays for Joseph Agassi on the Occasion of His 90th Birthday*. Dodrecht: Springer.

Bartlett, S. J. (Ed.). (1992). *Reflexivity. A Source-Book in Self-Reference*. Amsterdam: North-Holland.

Bartlett, S. J., Suber, P. (Eds.) (1987). *Self-Reference. Reflections on Reflexivity*. The Hague: Nijhoff.

Barwise, J., Etchemendy, J. (1987). *The Liar: An Essay on Truth and Circularity*. Oxford: Oxford University Press.

Beall, J. C., Armour-Garb (Ed.). (2009). *Deflationism and Paradox*. Oxford: Clarendon Press.

Black, M. (1949). *Language and Philosophy*. Ithaca: Cornell University Press.

Bolander, T., Hendricks, V. F., Pedersen, S. A. (Eds.) (2006). *Self-Reference*. Stanford: CSLI Publications.

Church, A. (1976). Comparison of Russell's resolution of the semantical antinomies with that of Tarski. *The Journal of Symbolic Logic*, 41, 747–760; repr in Martin (1984), 291–306.

Church, A. (2019). *Collected Papers*. The MIT Press.

Cieśliński, C. (2017). *The Epistemic Lightness of Truth. Deflationism and its Logic*. Cambridge: Cambridge University Press.

Cook, R. T. (2014). *The Yablo Paradox. An Essay on Circularity*. Oxford: Oxford University Press.

Davidson, D. (1967). Truth and meaning. Synthese, 17, 304–323; repr in Davidson (1984), 37–54.

Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.

De Fioro, C. (2013). *La forma della verità. Logica e filosofia nell'opera di Alfred Tarski*. Milano: Mimesis.

Fenstad, J. E. (2004). Tarski, truth, and natural language. In Z. Adamowicz, S. Artemov, D. Niwiński, E. Orłowska, A. Romanowska, J. Woleński (Eds.), 18–26.

Gabbay, D. M., Guenther F. (Eds.) (2011) *Handbook of Philosophical Logic*, v. 11, Kluwer: Dordrecht.

Gruber, M. (2016). *Alfred Tarski and the "Concept of truth in formalized languages". A Running Commentary with Consideration of the Polish Original and the German Translation*. Dordrecht: Springer.

Halbach, V. (2011). *Axiomatic Theories of Truth*. Cambridge: Cambridge University Press.

Horsten, L. (2011). *The Tarskian Turn. Deflationism and Axiomatic Truth*. Cambridge, Mass.: The MIT Press.

Kashtan, D. (2017). An observation about Truth (with Implications for Meaning and Language). Jerusalem: The Hebrew University (unpublished PhD dissertation).

Kokoszyńska, M. (1936). W sprawie względności i bezwzględności prawdy (Concerning relativity and absoluteness of truth). *Przegląd Filozoficzny*, 39, 424–425.

Kripke, S. (1975). An outline of a theory of truth. *The Journal of Philosophy*, 72, 690–716; repr. in Kripke (2011), 75–98.

Kripke, S. (2011). *Philosophical Troubles. Collected papers*, v. 1. New York: Oxford University Press.

Kripke, S. (2019). Philosophical Troubles.*Collected papers*, v. 2. New York: Oxford University Press.

Kripke, S. (2019a). Ungroundedness in Tarskian languages. In Kripke (Ed.).

Kuźniar, A., Odrowąż–Sypniewska, J. (Eds.) (2016). *Uncovering Facts and Values. Studies in Contemporary Epistemology and Political Philosophy*. Leiden: Brill/Rodoppi

Łukasiewicz, J. (1910). *O zasadzie sprzeczności u Arystotelesa* (On the principle of contradiction in Aristotle). Kraków: Wydawnictwo Polskiej Akademii Umiejętności; Eng. tr. (by H. R. Heine) *The Principle of Contradiction in Aristotle: A Critical Study*. North Charleston: CreateSpace Independent Publishing: North Charleston, SC 2018.

Łukasiewicz, J. (1915). O nauce (On science). In *Poradnik dla Samouków*. Warszawa: Heflich i Michalski, XV–XXXIX.

Martin, R. L. (Ed.). (1984). *Recent Essays on Truth and the Liar Paradox*. Oxford: Clarendon Press.

Maudlin, T. (2004). *Truth and Paradox. Solving the Riddles*. Oxford: Clarendon Press.

McGee, V. (1991). *Truth, Vagueness and Paradox*. Indianapolis: Hackett Publishing Company.

Patterson, D. E. (2012). *Alfred Tarski. Philosophy of Language and Logic*. Basingstoke: Palgrave Macmillan.

Priest, G. (2006). *Doubt Truth to Be a Liar*. Oxford: Clarendon Press.

Putnam, H. (1975). *Mind, Language and Reality. Philosophical Papers*, v. 2. Cambridge: Camrbidge University Press

Putnam, H. (1975a). Do true assertions correspond to reality? In Putnam (1975), 70–84.

Rüstow, A. (1908). *Der Lügner*. Leipzig: Teubner.

Simmons, K. (1993). *Universality and the Liar. An Essay on Truth and the Diagonal Argument*. Cambridge: Cambridge University Press.

Smorynski, C. (1985). *Self-Reference and Modal Logic*. Berlin: Springer.

Soames, S. (1999). *Understanding Truth*. Oxford: Oxford University Press.

Stegmüller, W. (1957). *Das Wahrheitsbegriff und die Idee der Semantik*. Wien: Springer.

Tarski, A. (1930–1931). O pojęciu prawdy w odniesieniu do sformalizowanych nauk dedukcyjnych (On the concept of truth in reference to formalized deductive sciences). *Ruch Filozoficzny*, 12, 210–211; repr. in Tarski (1986), v. 4, 559.

Tarski, A. (1930). Über definierbare Zahlenmengen. *Annales de Société Polonais Mathématique*, 9, 206–207; repr. in Tarski (1986), v. 4, 560–561.

Tarski A. (1931). Sur les ensembles définissables de nombres réels. I. *Fundamenta Mathematicae*, XVII, 210–239; Eng. tr. in Tarski (1956), 110–142.

Tarski, A. (1932). Der Wahrheitsbegriff in der Sprachen der deduktiven Wissenschaften. *Akademie der Wissenschaften in Wien. Matematisch-Wissenschaftliche Anzeiger*, 69, 23–25; repr. in Tarski (1986), v. 1, 613–617.

Tarski, A. (1933). *Pojęcie prawdy w językach nauk dedukcyjnych* (The Concept of Truth inLanguages of Deductive Sciences). Warszawa: Towarzystwo Naukowe Warszawskie; Germ. tr. (with additions) as Tarski (1935).

Tarski, A. (1935). Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica*, 1, 261–405; repr. in Tarski (1986), v. 2, 51–198. Engl. tr. *The Concept of Truth in Formalized Languages*, in Tarski (1956), 152–278 (page-references to Tarski 1933).

Tarski A. (1936). Grundlegung der wissenschaftliche Semantik. In *Actes du Congès International de Philosophie Scientifique*, v. 3, 1–8. Paris: Herman; Eng. tr. in Tarski (1956), 401–408.

Tarski A. (1936a). Über den Begriff der logischen Folgerung. In Actes du Congès International de Philosophie Scientifique, v. 7. Paris: Herman, 1–11; Eng. tr. in Tarski (1956), 409–420.

Tarski, A. (1936b). [Głos w dyskusji nad referatem M. Kokoszyńskiej] (A Contribution to the Discussion of M. Kokoszyńska's Paper). Przegląd Filozoficzny 39, 425; repr. in Tarski (1986), v. 4, 701).

Tarski, A. (1944). Semantic theory of truth and the foundations of semantics. *Philosophy and Phenomenological Research* 4, 341–395; repr. in Tarski (1986), v. 2, 661–697.

Tarski, A. (1956). *Logic, Semantics, Metamathematics. Papers of 1923 to 1938* (tr. by J. H. Woodger). Oxford: Clarendon Press; 2nd ed., Indianapolis: Hackett Publishing Company 1983.

Tarski, A. (1969). Truth and proof. *Scientific American*, 220(6), 63–77; repr. in Tarski (1986), v. 4, 399–423.

Tarski, A. (1986), *Collected Papers*, v. 1–4. Basel: Birkhäuser; 2nd ed. 2018.

Tarski, A., Givant, S. (1987). *A Formalization of Set Theory without Variables*. Providence, R. I.: American Mathematical Society.

Twardowski, K. (1997). *Dzienniki* (Daily Diaries), v. 1–2. Toruń: Adam Marszałek.

Visser, A. (2011). Semantics and the liar paradox. In Gabbay, Guenther (2011), 149–240.

Wang, H. (1986). *Beyond Analytic Philosophy. Doing Justice to What We Know*. Cambridge, MA: The MIT Press.

Woleński, J. (1995). Logic and falsity. *Logica 1994*, 95–105.

Woleński, J. (2016). The paradox of analyticity and related issues. In Abeles, Fuller (2016), 135–138; repr. in Woleński (2018), 205–208.

Woleński, J. (2016b). Formal and informal aspects of the semantic theory of truth. In Kuźniar, Odrowąż–Sypniewska (2016), 55–66; repr. in Woleński (2018), 195–204.

Woleński, J. (2017a). The story of the German translation of Tarski's book of truth. In Bar-Am, Gattei (Eds.), 81–90.

Woleński, J. (2018). *Logic and Its Philosophy*. Frankfurt a. M.: Peter Lang.

Woleński, J., Köhler, E. (Eds.) (1999). *Alfred Tarski and the Vienna Circle. Austro-Polish Connection in Logical Empiricism*. Dordrecht: Kluwer

Yablo, S. (1993). Paradox without self-reference. *Analysis*, 33, 251–252.

Yaqūb, A. M. (1993). *The Liar Speaks the Truth. A Defence of the Revision Theory of Truth*. Oxford: Oxford University Press.

Chapter 8
Semantic Theory of Truth—Formal Aspects

Abstract This chapter contains a detailed account of **STT** as a formal theory. The exposition considers truth as truth in a model. Firstly, truth-definition as satisfaction by all sequences of objects is explained. Arithmetic of natural numbers and its models play the crucial role in presenting various results concerning the concept of truth, particularly limitative theorems and the undefinability of arithmetical truth in arithmetic itself. Models constructed on terms are used as tools for defining the denotations of sentences in models. The last section reports Gödel's and Tarski's views on limitative theorems and truth.

8.1 Introduction

As it follows from numerous earlier remarks, the borderline between formal and informal aspects of logical (in the wide sense) constructions cannot be drawn in a precise manner. Chapter 7, "Semantic Theory of Truth—Informal Aspects", contains several assertions also concerning formal matters. Conversely, this chapter will also go into some informal aspects. The content of this chapter breaks down as follows. Section 8.2 present **SDT** in a rigorous, mathematical way and describes some consequences of this definition. The next four sections explore metalogical (metamathematical) issues as related to the concept of truth from the arithmetical point of view. Section 8.7 describes models constructed on terms and their application to the problem of denotations of sentences. The last section reports and compares Gödel's and Tarski's views pertain to the undefinability of truth. In many places of the present chapter, I follow Grzegorczyk 1974 (see also Gómez-Torrente 2004). All considerations assume that **L** is first-order and consists of denumerably (infinitely) many open formulas and/or sentences. That does not mean that other languages are entirely outside the scope of the outlined construction, but I do not get into this issue—except for some marginal remarks.

J. Woleński, *Semantics and Truth*, Logic, Epistemology, and the Unity of Science 45, https://doi.org/10.1007/978-3-030-24536-8_8

8.2 SDT as a Formal Construction

The formulation of (20) in Chap. 7 is provisional, because it does not correspond to the full definition of satisfaction. The earlier explanations concerned the simplest case, namely satisfaction of monadic open formulas, that is, of the form $P(x)$. What about the formula (a) 'x is a larger city than y', which expresses the relation of being a larger city? The cities New York and Chicago, taken in this sequence satisfy (a), but not in the reverse one. Since relations are sets of ordered pairs, we can say that (a) is satisfied by the ordered pair <New York, Chicago>, but not by the ordered pair <Chicago, New York>. Since formulas can have arbitrary length, we need a generalization of this procedure in order to have a uniform way of dealing with all cases. This was Tarski's motivation for introducing the concept of satisfaction by means of infinite sequences of objects. Since formulas are of arbitrary but always finite length, infinite sequences have a sufficient number of members to cover the satisfaction of all possible cases of particular formulas. Thus, the canonical articulation is as follows (at the moment this explanation pertains to open formulas:

(1) A is satisfied by an infinite sequence $\mathbf{s} = \langle s_1, s_2, s_3, \ldots \rangle$, where s_n ($n \geq 1$) refers to the *nth* term of **s**.

Sequences in the sense of (1) help in formulating the official definition of satisfaction in the following way (I use terminology from Chap. 5, Sect. 5.2.4, but I simplify indexing and restrict terms to individual variables and individual constants)

(Df1) (**a**) 'P_j (t_1, …., t_k)' $\in$ $\mathbf{SAT}(\mathbf{s}, \mathfrak{I}) \Leftrightarrow \langle \mathfrak{I}$ ('t_1'), …, $\mathfrak{I}$('t_k')$\rangle \in \mathbf{R_j}$ (=$\mathfrak{I}$ ('P_j');
(**b**) '$\neg A$' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I}) \Leftrightarrow$ 'A' $\notin \mathbf{SAT}(\mathbf{s}, \mathfrak{I})$;
(**c**) '$A \wedge B$' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I}) \Leftrightarrow$ 'A' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I})$ and 'B' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I})$;
(**d**) '$A \vee B$' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I}) \Leftrightarrow$ 'A' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I})$ or 'B' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I})$;
(**e**) '$A \Rightarrow B$' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I}) \Leftrightarrow$ '$\neg A$' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I})$ 'B' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I})$;
(**f**) '$A \Leftrightarrow B$' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I}) \Leftrightarrow$ '$A \Rightarrow B$' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I})$ and '$B \Rightarrow A$' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I})$;
(**g**) '$\forall x_i A(x_i)$' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I}) \Leftrightarrow$ '$A(x_i)$' $\in \mathbf{SAT}(\mathbf{s}', \mathfrak{I})$, for every sequence $\mathbf{s}'$, which differs from the sequence **s** at most at the ith place;
(**h**) '$\exists x_i A(x_i)$' $\in \mathbf{SAT}(\mathbf{s}, \mathfrak{I}) \Leftrightarrow$ '$A(x_i)$' $\in \mathbf{SAT}(\mathbf{s}', \mathfrak{I})$, for some sequence $\mathbf{s}'$, which differs from the sequence **s** at most at the ith place.

The first clause establishes the satisfaction-conditions for atomic formulas that refer to relations (sets can be considered as one-placed relations). Conditions (**b**)–(**f**) repeat the semantic definitions of propositional connectives given in Chap. 5, Sect. 5.2.2, (**g**) and (**h**) concern quantifiers and say that an (open) universal formula is satisfied by every sequence, but an existential formula by some sequence ('differs at most at most ith place' is a technical phrase to capture the intended meaning). The reference to an interpretation $\mathfrak{I}$ indicates its role in correlation of expressions and their references, for instance predicates and relations. Since $\mathfrak{I}$ is always associated with a model **M**, the expression 'A' ∈ **SAT**(**s**, $\mathfrak{I}$) can be replaced by the phrase 'A' ∈ **SAT**(**s**, **M**) (a formula A is satisfied by a sequence **s** in a model **M**).

The construction outlined above has obvious technical merits, but it does raise some doubts as to its literal content, because we are not quite entitled to say that formulas are satisfied by sequences of objects, because sequences are functions from positive integers to arbitrary non-empty sets. Formally speaking, we have the mapping **s**: **N** → **Y**, where **N** is the set of natural numbers and **Y** is an arbitrary non-empty set; if **X** is a finite set, then **s** is a finite sequence. Sequences then consist of ordered pairs of the type $\boldsymbol{s_1}$ = <**1**, $\boldsymbol{y_1}$>, $\boldsymbol{s_2}$= <**2**, $\boldsymbol{y_2}$>, etc., where $\boldsymbol{y}_k$ ($1 \leq k$) ∈ **X**. Consequently, formulas are satisfied not by individual objects, but by pairs consisting of positive integers and objects (from **X**) mapped by **s**. In particular, if we say that Warsaw satisfies the formula 'x is the capital of Poland', we should say that this formula is satisfied by the object <Warsaw>, that is, by the one-termed sequence with Warsaw as its sole term.

(**DG1**) I use the following notational conventions. Variables are represented by normal letters, indexed or not (the second case concerns examples), sequences by the letters **s**, **s′**, …, but objects (including the terms of sequences) by the bold letters, for instance, $\boldsymbol{a_1}$, $\boldsymbol{a_2}$, …; $\boldsymbol{s_1}$, $\boldsymbol{s_2}$, …. The notation $\boldsymbol{x_n}$ ∈ **X** means that the object $\boldsymbol{x_n}$ belongs to the set **X**; the context x_n ∈ **Var** means 'a variable x_n belongs to the set **Var** of variables'.►

Yet we can says that formulas are satisfied or not by sequences (in an intuitive sense) of objects in a very simple way. If **s** is a function from **N** to **X**, then **N** is its domain, but **X**—its range. We construct an image **i**(**N**) of the set **N** given by **s**: **N** → **X**, that is, the set of all elements of **X** that are values of the function **s.** Formally, **x** ∈ **i**(**N**) ⇔ **x** = **i**($\boldsymbol{n}$)). In order to apply this construction to our problem, we must first clarify the details concerning the concept of sequence to be used when we speak about satisfaction of formulas. Clearly, it is not enough to say that such sequences are functions (or mappings) from **N** to **X**. Sequences used in formal semantics (sequences of objects) have something to do with variables and interpretations. Now we can rewrite (**Df1a**) as

(2) '$P_j(x_1,\ldots,x_k)$' is satisfied in the model **M** by the sequence **s** = <$\boldsymbol{s_1}$, …, $\boldsymbol{s_2}$,…, $\boldsymbol{s}_k$, …> given by an image **i**(**N**) ⇔ <$\boldsymbol{s}_n$, …, $\boldsymbol{s_k}$> ∈ $\mathbf{R_j}$, where $\mathbf{R_j}$ = **v**(P_j).

How to construct a required image given by the $\boldsymbol{s}$ used in (2)? Let **Var** be the set of individual variables of **L.** As usually, we assume that **Var** is denumerably infinite. Then we define a mapping **s**: **N** → **Var**, and the image **i**(**N**) of **s**. This step generates an ordering of variables by indices that is the sequence x_1, x_2, x_3, …. The next move consists in using a mapping from indexes (the set **Ind**) representing variables to objects in **U** (**U**—because we are working with a given model). That gives a function **s′**: **Ind** → **U**, which is in fact the function from **N** to **U** (**s** = **s′**). Further, we define the image of **Ind** to objects and obtain the set of objects corresponding to individual variables. Finally, we say that formulas are satisfied by images of the set of variables given by sequences. Since images map variables into objects, we can say that formulas are satisfied (or not) by semantic images of variables, that is, objects associated with variables ordered by $\mathbf{s}$; formally speaking, objects as semantic images are the results of composing the functions $\mathbf{s}'$ and $\mathbf{s}$.

One may ask whether there is any difference between satisfaction and interpretation, because semantic images satisfy formulas and interpret variables. Yes, there is a difference. If we interpret variables in a way, there is no sense to ask whether variables are interpreted or not because for are by the definition of valuation. Yet particular valuations (interpretations) lead to satisfaction of formulas or not, because that also depends on the denotations of predicates. Every predicate is associated with a subset of **U** or its *n*-termed Cartesian product (a relation). Since values of constants are fixed, while interpretations of variables vary from one sequence to another, some valuations result in satisfaction, whereas others do not. For example, if Cracow values the variable *x* in '*x* is the capital of Poland', the related semantic image does not satisfy this formula, but if we take Warsaw as the image, it becomes satisfied, although both valuations interpret the variable *x*. This way of thinking about satisfaction shows why satisfaction of open formulas does in fact depend on free variables. Thus, we can restrict infinite sequences to those of their sub-sequences that have only terms corresponding to free variables. This possibility concerns sequences longer than those determined by the actual number of free variables in the formulas under consideration.

(**DG2**) If $\mathbf{s}$ is an infinite sequence and *A* has *n* free variables, only *n* terms of $\mathbf{s}$ are relevant to *A's* being satisfied or not. Hence, another possibility (see Tarski 1933, p. 195, Popper 1955, p. 337) to define the satisfaction relation is to introduce sequences of a sufficient finite length.►

What about sentences? Consider the example with New York and Chicago, but starting with the formula.

(3) x_1 is a larger city than x_2.

This formula is satisfied by every ordered pair $\langle \boldsymbol{s_1}, \boldsymbol{s_2} \rangle$ such that $\boldsymbol{s_1} = \mathbf{v}(x_1)$ and $\boldsymbol{s_2} = \mathbf{v}(x_1)$ are cities, and $\boldsymbol{s_1}$ is larger than $\boldsymbol{s_2}$. In particular, the pair <New York, Chicago> (I do not use bold characters, but this causes no complications from the intuitive point of view) satisfies (3). Note that the sequence $\langle \boldsymbol{s_1}, \boldsymbol{s_2} \rangle$ can be enlarged by adding an arbitrary number of terms in order to have an infinite sequence $\langle \boldsymbol{s_1}, \boldsymbol{s_2}, \boldsymbol{s_3}, \ldots, \boldsymbol{s_k}, \ldots \rangle$, but this operation is irrelevant to satisfaction. Informally speaking, if a sequence $\langle \boldsymbol{s_1}, \boldsymbol{s_2} \rangle$ satisfies (or not) the formula (3), the same applies to the sequence $\langle \boldsymbol{s_1}, \boldsymbol{s_2}, \boldsymbol{s_3}, \ldots, \boldsymbol{s_k}, \ldots \rangle$, because the terms $\boldsymbol{s_1}$, $\boldsymbol{s_2}$ are the only one that are significant for the satisfaction business related to (3). Now substitute Chicago for x_2. That gives

(4) x_1 is a larger city than Chicago.

This formula is satisfied by the sequence $\langle \boldsymbol{s_1} \rangle$ such that $\boldsymbol{s_1} = \mathbf{v}(x_1)$, is a city and $\boldsymbol{s_1}$ is larger than Chicago, in particular by the object <New York>. Enlarging the sequence <New York> by adding an arbitrary number of terms does not change the situation. Every sequence of the form <New York, $\boldsymbol{s_2}, \boldsymbol{s_3}, \ldots, \boldsymbol{s_k}, \ldots$> satisfies the formula (4).

Finally, consider

(5) New York is a larger city than Chicago.

This expression is just a sentence, not an open formula. Since it has no free variables, its satisfaction does not depend on valuations of free variables. Hence, every infinite sequence of the form $<s_1, s_2, s_3, \ldots, s_k, \ldots>$ satisfies (5). In other words, we can replace s_k by an arbitrary object (remember that s_k is the value of the variable x_k) and this step has no relevance for the satisfaction of (5). It is satisfied because New York is a larger city than Chicago. Another way to the same result consists in using a theorem of **FOL** that if *A* is a sentence $\forall x_i A \Leftrightarrow A$. Assume that a sequence **s** satisfies (5). By clause (**Df1 g**), formula *A* is also satisfied by every sequence **s′** which differs from **s** at most at the *i*th place. Since *A* has no free variables, the *i*th place can be arbitrarily chosen from terms of **s′**. This means, that every sequence satisfies *A*. This reasoning implies that if a sentence *A* is satisfied by at least one sequence, it is also satisfied by any other sequence. Thus, we obtain the following statements

(6) A sentence is satisfied by all sequences if and only if it is satisfied by at least one sequence.
(7) A sentence is not satisfied by all sequences if and only if it is satisfied by no sequence.

Both assertions lead to

(8) If *A* is a sentence it is satisfied by all sequences or is satisfied by no sequence.

If (see Chap. 7, Sect. 7.5), we assume that truth is considered a special instance of the satisfaction relation, (6) and (7) lead to the following definition (Tarski 1933, p. 195):

(**Df2**) (**a**) '*A*' is true in **M** if and only if '*A*' is satisfied by every infinite sequence of objects from the universe of **M** (equivalently: by at least one such sequence);
(**b**) '*A*' is false in **M** if and only if '*A*' is not satisfied by some infinite sequence of objects from the universe of **M**.

Due to (**Df2a)**, condition (**Df2b**) is equivalent to the statement that '*A*' is false if and only if it is satisfied by no sequence.

(**DG3**) If **M** is a canonical model—that is, such that every object in **U** has a name (see Chap. 5, Sect. 5.2.4)—the clause (**Df1a**) becomes simpler, because it is enough to works with values of individual constants. This circumstance simplifies the entire (**Df1)**, since its first condition formulates the inductive basis. However, it is the special case only.►

Returning once again to the sequence $\mathbf{s} = <s_1, s_2, s_3, \ldots, s_k, \ldots>$, assume that a given formula *A* has *k* free variables. Thus, we can consider the sequence $\mathbf{s'} = <s_1, \ldots, s_k>$ as relevant for the satisfaction of *A* (**s′** is a sub-sequence of **s**). Suppose that free variables are eliminated by substituting individual constants for them, or by quantification. Intuitively, elimination of a free variable x_i ($1 \leq i \leq k$) permits deleting it from the list of free variables, and the same applies to the term s_i. Terms

s_{k+1}, s_{k+2},... of **s** are the only erasable after determining s′. When the deletion-process reaches the term s_1, the object <> becomes the remainder of **s′** as well as **s**. Call it the empty sequence. Using this concept leads to the third truth-definition, which is equivalent to (**Df**2):

(**Df3**) (**a**) '*A*' is true if and only '*A*' is satisfied by the sequence <>.
(**b**) '*A*' is false if and only if '*A*' is not satisfied by the sequence <>

(**DG4**) The definition (**Df3**) was also given by Tarski (Tarski 1933, p. 195). He remarked on the occasion:

> Regarding the concept of truth, it is to be noted that—according to the above treatment—only *one* sequence, namely 'the empty' sequence which has no members at all, can satisfy a sentence, i.e. the function without free variables; we should then have to call those sentences true which are actually satisfied by the 'empty' sequence. A certain artificiality attached to this definition will doubtless displease all those who are not sufficiently familiar with the specific procedures which are commonly used in mathematical constructions.

Tarski himself preferred (**Df2a**) in the version with infinite sequences. It seems that he considered it as the most plausible from the intuitive point of view. This view has some justification. On the other hand, one might remark that, due to the noted equivalences in defining truth, either all versions are somehow intuitive or none of them. I will return to this issue in Chap. 9, Sect. 9.2.►

Introducing the object < > and calling it the empty sequence requires a generalization of the concept of sequence. The usual definition of sequences assumes that they are mappings from natural numbers to non-empty sets. Presumably, the 'empty' sequence is a mapping from $\varnothing \rightarrow \varnothing$, that is, the empty function. Barwise 1975, p. 83 defines the truth of a formula by the empty function, but he does not correctly call it a sequence, because the 'empty' sequence is not a sequence under the official definition, and it was probably Tarski's reason to write the 'empty' sequence, not simply—the empty sequence. Feferman 1989, p. 124 introduces the concept of the empty (without quotes) sequence as 0-termed, that is, precisely the mapping $\varnothing \rightarrow \varnothing$. Thus, we can say that **s** is a sequence if it is the empty function—or a mapping from **N** to **X**, provided **X** is not empty. Yet another approach is suggested in Gries, Schneider 1993, p. 251 (see also Woleński 2003). They construct an axiomatic theory of finite sequences. The axioms establish when an object belongs to the set **Seq** of sequences. The theory is based on the concept of the empty sequence as the sole primitive idea. The axioms are as follows (the symbol ∇ refers to the operation of adding a new element; we also assume that sequences **s**, **t** are defined over a fixed set **X**):

(A1) $<> \in \mathbf{Seq}$
(A2) $\boldsymbol{x_k} \nabla \boldsymbol{s} \in \mathbf{Seq}$
(A3) $\boldsymbol{x_n} \nabla \boldsymbol{s} \neq <>$
(A4) $\boldsymbol{x_k} \nabla \boldsymbol{s} = \boldsymbol{x_n} \nabla \boldsymbol{t} \Leftrightarrow \boldsymbol{x_k} = \boldsymbol{x_n} \wedge \boldsymbol{s} = \boldsymbol{t}$

Since truth is a special case of satisfaction, (**Df1**) can be transformed into a truth-definition. I use <> as the device to do it, but the same can be accomplished via infinite sequences. Moreover, **M** replaces **s**, $\mathfrak{I}$ (an interpretation is always associated with every model), and **v** (‘t_1’), …, **v**(‘t_k’) is used instead <$\mathfrak{I}$ (‘t_1’), …, $\mathfrak{I}$(‘t_k’)> ; **VER**(**M**) (it replaces **SAT**(**M**) refers to the set of truths in **M**. With these conventions, we obtain

(Df4) (**a**) ‘P_j (t_1, …., $t_{k)}$’ ∈ **VER**(**M**) ⇔ <**v** (‘t_1’), …, **v**(‘t_k’)> ∈ $\mathbf{R_j}$;
(**b**) ‘¬A’ ∈ **VER**(**M**) ⇔ ‘A’ ∉ **VER**(**M**);
(**c**) ‘$A \wedge B$’ ∈ **VER**(**M**) ⇔ ‘A’ ∈ **VER**(**M**) and ‘B’ ∈ **VER**(**M**);
(**d**) ‘$A \vee B$’ ∈ **VER**(**M**) ⇔ ‘A’ ∈ **VER**(**M**) or ‘B’ ∈ **VER**(**M**);
(**e**) ‘$A \Rightarrow B$’∈ **VER**(**M**) ⇔ ‘¬A’ ∈ ‘B’ or B ∈ **VER**(**M**);
(**f**) ‘$A \Leftrightarrow B$’∈ **VER**(**M**) ⇔ ‘$A \Rightarrow B$’ ∈ **VER**(**M**) and ‘$B \Rightarrow A$’ ∈ **VER**(**M**);
(**g**) ‘∀ $x_iA(x_i)$’ ∈ **VER**(**M**) ⇔ ‘$A(x_i)$’ is satisfied by <> ;
(**h**) ‘$\exists x_iA(x_i)$’ ∈ **VER**(**M**) ⇔‘$\neg\forall x_i\ \neg A(x_i)$’ ∈ **VER**(**M**).

Let us look at the consequences of **SDT** in the above formulation. Since it assumes resources to meet **LP** and similar paradoxes, its consistency against semantic antinomies is guaranteed. As I already noted in Chap. 7, Sect. 7.4, the issue of circularity, etc. is not simple. According to Tarski, **SDT** is formulated in the morphology of **ML**. However, inspection of (**Df2**)–(**Df4**) immediately leads to the conclusion that set theory is essentially employed in using sequences, sets, etc. Even though the set-theoretical apparatus is very elementary, it exceeds pure logic, at least according to contemporary views. Hence, when Tarski says that the morphology of **ML,** as the framework of **SDT**, is reducible to logic, his view was perhaps justified in the 1930s, but must be corrected in our times. We may well say (see Kokoszyńska 1936) that so-called extended syntax is at work here, but this requires an addition that extended syntax comprises an amount of set theory (see Casari 2006 for mathematics employed in **STT**). On the other hand, to confirm what I said at the end of Chap. 7, the formulations (**Df2**)–(**Df4**) do not employ the concept of truth, just as (**Df1**) does not define satisfaction in a circular manner. One can say that **SDT** proceeds as a typical mathematical construction based on a portion of set theory. Although some philosophers—for instance, Husserl and his followers—will probably be dissatisfied by this situation vis-a-vis their claim that philosophical constructions have to be free of presuppositions, the defenders of **SDT** (and similar constructions) can reply that (a) conformity to mathematical practice is more important than established a priori metaphilosophical postulates, and that (b) an informal understanding of **ML** is inevitable for logical constructions pertaining to **L**.

Condition (b) of **CT** (see Chap. 7, Sect. 7.5) is satisfied because if A ∈ **VER**, then A ∈ **L**. The clause (**Df4a**) establishes the translation of ‘P^j(t_1, …., t_k)’ into **ML**. Since ‘t_1’, …., ‘t_k’ are individual constants, their values are concrete objects from **M**. This observation leads to rewriting (**Df4a**) as

(9) ‘P_j (t_1, …., t_k)’ ∈ **VER**(**M**) ⇔ <$\boldsymbol{a}_\mathbf{1}$, …, $\boldsymbol{a}_k$> ∈ $\mathbf{R_j}$.

Now (8) implies

(10) 'P_j (t_1, …., $t_{k)}$' $\in$ **VER**(**M**) $\Leftrightarrow$ P_j (t_1, …., t_k)*,

which is a special instance of **T**-scheme. Since (**Df4a**) is the inductive clause of the entire definition (**Df4**), the same applies (8) and (9). Thus, we have

(11) (a) 'P_j (t_1, …., t_k)' $\in$ **VER**(**M**) $\Leftrightarrow$ P_j (t_1, …., t_k)*;

(b) '$\neg A$' $\in$ **VER**(**M**) $\Leftrightarrow$ ($\neg A$)*;

(c) '$A \wedge B$' $\in$ **VER**(**M**) $\Leftrightarrow$ ($A \wedge B$)*;

(d) '$A \vee B$' $\in$ **VER**(**M**) $\Leftrightarrow$ ($A \vee B$)*;

(e) '$A \Rightarrow B$'$\in$ **VER**(**M**) $\Leftrightarrow$ ($A \Rightarrow B$)*;

(f) '$A \Leftrightarrow B$'$\in$ **VER**(**M**) $\Leftrightarrow$ ($A \Leftrightarrow B$)*;

(g) '$\forall x_i A(x_i)$' $\in$ **VER**(**M**) $\Leftrightarrow$ ($\forall x_i A(x_i)$*;

(h) '$\exists x_i A(x_i)$' $\in$ **VER**(**M**) $\Leftrightarrow$ ($\exists x_i A(x_i)$*.

This reasoning completes the justification that **SDT** satisfies **CT**. In particular, every the definition in question entails **T**-equivalence for all sentences of **L**.

The next matter concerns (**BI**). A possible proof runs as follows (see Grzegorczyk 1974, pp. 287–288; I omit in-quoting by ''). Assume that $A \notin$ **VER** (**M**), where A is a sentence. According to (**Df2**), this means that not every a sequence satisfies A. By **FOL**, it entails that there exists a sequence **s** such that **s** does not satisfy A. Furthermore, by (**Df1b**), **s** satisfies $\neg A$. Since A is a sentence, so is $\neg A$. By (**Df2a**), $\neg A$ is true. We proved

(12) $A \notin$ **VER**(**M**) $\Rightarrow$ $\neg A \in$ **VER**(**M**).

The proof of (11) is similar. Having (11) and its converse, the theorem

(13) $A \notin$ **VER**(**M**) $\Leftrightarrow$ $\neg A \in$ **VER**(**M**),

which can be interpreted as

(14) For every A and **M**, every sentence is either true in **M** or false in **M**.

In fact, transposing of (13) results in

(15) $\neg A \notin$ **VER**(**M**) $\Leftrightarrow$ $A \in$ **VER**(**M**).

Decomposing (15) gives

(16) $\neg A \notin$ **VER**(**M**) $\wedge$ $\neg A \notin$ **VER**(**M**) $\vee$ $A \in$ **VER**(**M**) $\wedge$ $\neg A \notin$ **VER**(**M**).

The first conjunction in (16) can be dropped as inconsistent. The second possibility consists in truth of A and falsehood of $\neg A$. It implies that either A is true or A is false. This conclusion is equivalent to (**BI**). The above demonstration of the principle of bivalence uses classical logic. In fact, steps leading to (12) employ the De Morgan laws for quantifiers (passing from 'not every sequence satisfies' to 'there exists a sequence which does not satisfy') and the analysis of (15) is based on De

Morgan laws in **PC** concerning denials of disjunctions and conjunctions. Because these laws are characteristic for classical logic, one might complain that the given justification of (**BI**) suffers from circularity—since classical logic is used in order to obtain its basic metalogical principle.

Although it is not without importance that **SDT** entails (**BI**), even if with the help of classical logic, the situation can be improved. First of all, (**BI**) can be directly derived from (8) via the definition (**Df2**) via using <>. Although the empty sequence is somehow strange, it is a well-defined and concrete object. Thus, we can express some facts about truth constructively, for instance, the definition of falsehood. In particular, (**Df3b**) says that A is false if and only if A is not satisfied by <>. With the empty sequence we can prove bivalence more constructively, that is, without appeal to critical theorems of classical logic, De Morgan laws for **PC** and **FOL**, in particular. This is philosophically important because it shows that bivalence holds in **ML** without assuming all of classical logic in the metatheory of **SDT**. That is not to suggest that the semantic theory of truth is thereby constructive. It is not, because the concept of model is not constructive in general (this question will be discussed more extensively in Sect. 8.6.

The set **VER**(**M**) has some intuitively expected metalogical properties. They are summarized (I use the notation introduced in Chap. 5, Sect. 5.2.4; the list below is incomplete owing to the vagueness of 'intuitively expected') as follows;

(17) (a) **VER**((**M**) $\in$ **CONS**;
(b) **VER**(**M**) $\in$ **SYS** (Cn**VER**(**M**) $\subseteq$ **VER**(**M**));
(c) **VER**(**M**) is maximally consistent. This means that if $A \notin$ **VER**((**M**), then **VER**((**M**) $\cup$ $\{A\}$) is inconsistent;
(d) **VER**(**M**) $\in$ $\mathbf{COM}_{\mathbf{SYN}}$;
(e) **VER**(**M**) $\in$ **COMP**.

The first property is evident, because **VER**(**M**) cannot contain A and $\neg A$. If **L** does not have the negation, **VER**(**M**) $\neq$ **L**. So **VER**(**M**) absolutely consistent as well as negation consistent. (17b) has its justification in the fact that logical consequences of truths are true. The Lindenbaum lemma says that every consistent set of sentences has maximally consistent extension. If **X** is an arbitrary set of truth, **VER**(**M**) is its maximal enlargement. This justifies (17c). The next property follows directly from the Lindenbaum property of **VER**(**M**). The compactness of **VER**(**M**), that is, (17e), follows from (17a). Clearly, **VER**(**M**) is true if and only if its every finite subset is true. However, sets of true sentences are not generally axiomatizable (see Sect. 8.3). This fact makes the syntactic completeness of **VER**(**M**) is not as interesting as the case of axiomatizable sets of sentences. To supplement this survey, I note that (a) sets of the type **VER**(**M**) are not always decidable or, to put it more precisely—the most interesting sets of truths are just undecidable; (b) sets of the type **VER**(**M**) are not always axiomatizable, that is, they not always satisfy the condition **VER**(**M**) = Cn(**Ax**), where **Ax** $\subset$ **VER**(**M**).

By the semantic completeness theorem, every set of truths has a model, because it is consistent. However, this assertion does not imply that such sets have unique models. On the contrary (see Chap. 5, Sect. 5.3), the Löwenheim–Skolem property determines that first-order theories have many non-isomorphic models. What about the object **MOD**(**X**) such that for every **M**(**X**), **M**(**X**) $\in$ **MOD**(**X**) and **X** is a set of sentences of **L**? Informally speaking, **MOD** is a collection of models. A formal definition is provided by

(18) $\mathbf{M} \in \mathbf{MOD}(\mathbf{X}) \Leftrightarrow \forall A \in \mathbf{X}(A \in \mathbf{VER}(\mathbf{M}))$.

I did not use the term 'set' in introducing **MOD**. Given a set **X, MOD**(**X**) refers to all models of **X**. Assume that this collection is a set. Consider the set **Y** of all subsets of **MOD**(**X**). Every such subset is a set of models of **X**. By set theory, **Y** is larger than **MOD**(**X**). This contradicts the assumptions that **MOD**(**X**) contains all models of **X**. In fact, **MOD**(**X**) is a proper class (see Enderton 1972, p. 92; see also **DG12VII**). Although classes are subject to mathematical treatment, some limitations occur do here. For instance, the class of objects that satisfy a condition $A(x)$ exists provided this formula does not contain such quantifiers as 'for any class such that' or 'there is a class such that'. The condition $\forall A \in \mathbf{X}(A \in \mathbf{MOD}(\mathbf{X}))$ that occurs in (18) is predicative, since it does not quantify over classes. If **Y** is a class, it has a complement $-\mathbf{Y}$ which is also a class. Hence, if **MOD**(**X**) is a class, there is the class $-\mathbf{MOD}(\mathbf{X})$ to which belong structures in which at least one sentence A, such that $A \in \mathbf{X}$, is false. Similar considerations concern the object **VER**(**MOD**)—the set of all truths that hold in a class of models **M**. However, (17c) implies that **VER**(**M**) is a set. Two consequences follow from the last remarks. Firstly, we should expect limitations in defining collections of truths and models. Secondly, if we consider theories, which are typically a parts of all the truths in a given model, we can use more set-theoretical constructions than in the case of **VER**(**M**).

SDT involves a point that requires further analysis. If sentences are interpreted as true or false, perhaps they should have some references in interpretative structures. I am speaking not about models of false sentences. *Prima facie*, they cannot have references in models because they do not assert how thing are. More formally, we ask for what should be inserted in the place of [...] in the expression $\mathfrak{I}$('A', $\mathfrak{M}$) **= [...]. SDP** suggests nothing in this respect. To say that all infinite sequences, an arbitrary chosen sequence, or the sequence <> can be taken as the object in question is not satisfactory, because repeats the definition. We can eventually extend $\mathfrak{M}$ by adding two new objects, namely **1** (**Truth**) and **0** (**Falsehood**) as referents of sentences. This manoeuvre reveals Frege's ideas (see Chap. 3, Sect. 3.6), but without assuming that the concept of truth is not definable, and follows a possible semantics for **PC** using **1** and **0** as special objects. However, these new objects are artificial in **FOL,** because they are not set-theoretical constructs over **U.** One would like to see logical values as somehow associated with features of interpretations as structures. Consider sentences as zeroary predicates; if 'is P' is a monadic predicate, the sentences $P(a)$, $\forall x(P(x))$ and $\exists x(P(x))$ are zeroary predicates. Similarly, we can introduce zeroary properties and relations as denotations of sentences. The objects **1** and **0** are such, but that is only a stipulation only, not the result of analysis.

Another possible construction consists in introducing zeroary Cartesian powers of **U**. In general**,** $\mathbf{Y^0} = \{\varnothing\}$**,** for any set **Y** (see Bell, Machover 1977, p. 1). This convention is motivated by analogy with arithmetic in which $m^\circ = 1$, for any number m; in von Neumann's constructions of natural numbers $\{\varnothing\} = 1$. In general (see Poizat 2000, pp. 32–33), if **Y** is an arbitrary non-empty set, for instance a universe **U** of $\mathfrak{M}$, there are two zeroary functions defined on **U**, namely $\{\varnothing\}$ and $\varnothing$. Neither contains any elements from **U**; $\{\varnothing\}$ can be interpreted as truth, $\varnothing$—as falsehood. This construction motivates

(19) (a) $\mathfrak{I}$('*A*', $\mathfrak{M}$) = $\{\varnothing\}$ or $\varnothing$;
(b) $\mathfrak{I}$('*A*', $\mathfrak{M}$) = $\{\varnothing\}$, if '*A*' is true in the sense of (**SDT**);
(c) $\mathfrak{I}$('*A*', $\mathfrak{M}$) = $\varnothing$, if '*A*' false in the sense of (**SDT**).

At the first sight, (1) offers no natural semantic intuition. However, consider the equivalence

(20) '*A*' is satisfied by the sequence <> $\Leftrightarrow$ $\mathfrak{I}$('*A*', $\mathfrak{M}$) = $\{\varnothing\}$

as the starting point. Its left side says that there exists a sequence satisfying '*A*', but the right side tells us that the logical value of '*A*' is identified with the object $\{\varnothing\}$ as a subset of $\mathbf{U}^{\mathbf{n}}$. Since (20) is equivalent to

(21) '*A*' is not satisfied by the sequence <> $\Leftrightarrow$ $\mathfrak{I}$('*A*', $\mathfrak{M}$) = $\varnothing$,

we conclude that no subset of $\mathbf{U}^{\mathbf{n}}$ functions as the interpretation of false sentences (recall that **U** is not empty). This shows that assertions "'*A*' is not satisfied by <>' and '$\varnothing$ is the value of false sentences' should be carefully distinguished. Finally, tautologies are valued by $\{\varnothing\}$ in every interpretative structure, but contradictions—by $\varnothing$. Although $\{\varnothing\}$ and $\varnothing$ are artificial, they have the requisite metalogical properties. Finally, since semantic interpretations are inevitably associated with models, false sentences are falsehoods in models. This is an important conclusion, because it shows that both logical values are defined as related to **M**.

(**DG5**) Yet one would have such constructs that value sentences not only structurally (extensionally), but also intensionally. For reasons mentioned in (**DG5VI**), this is a difficult matter. In fact, we have $\mathfrak{I}$('*A*', $\mathfrak{M}$) = $\{\varnothing\}$ or $\mathfrak{I}$('*A*', $\mathfrak{M}$) = $\varnothing$, for any sentence of **L**. Consequently, this new proposal fully sanctions the fact that all true sentences have the same denotation and the same applies to all falsehoods. We can weaken slightly this consequence—regarded as quite unpalatable to many philosophers of logic—by introducing the notation $\mathfrak{I}$('*A*', $\mathfrak{M}$) = $\{\varnothing\}^{[\mathbf{A}^*]}$ and $\mathfrak{I}$('*A*', $\mathfrak{M}$) = $\varnothing\}^{[\mathbf{A}^*]}$ in order to indicate that logical values pertain to sentences translated in some particular way into **ML**. Due to the intensional parameter that was introduced, we can even propose a version of **T**-scheme as the formula $A \in$ **VER**(**M**) if and only if $\mathfrak{I}$('*A*', **M**) = $\{\varnothing\}^{[A^*]}$.

Another possible weakening (I do not say a solution, because a satisfactory intensional semantics for sentences is still an open issue) appeals to some ideas from model theory. Suppose that we have two languages **L** and **L′** such that the latter arose from the former by deleting some symbols. Technically, **L** is an extension of **L′** or **L′** is a restriction of **L.** Let **X** be a set of true sentences of **L** and **M**—its model. If $P(a) \in$ **X**, then $P(a) \in$ **VER**(**M**). Now we delete all extralogical symbol from **L** except '*P*' and '*a*'. This converts **L** into a language in which we can only expresses the sentence *P*(a) (I omit constructions obtained by adding propositional constants, for instance, $\neg P(a)$). This operation does not change the initial interpretation of the considered sentence, for example, that $\mathbf{v}(a) = \mathbf{u}$ and **v**

(P) = **P**. The structure $\mathbf{M}^{P(a)}$ = <{**u**}, **u**, **P**> in which the sentence $P(a)$ is true is the P(a)-reduct of **M**. In general, for every true atomic sentence A of **L**, we can define an A-reduct of **M**. The rest can be defined by induction; reducts of quantified formulas without free variables can be regarded as identical with reducts that satisfy open formulas in which no quantifier occurs. The reduction in question is conservative in the sense that it does not lead to any false sentence. On the other hand, expansion of **X** is not conservative. This set is consistent, but not necessarily syntactically complete. So there are sentences, say A, $\neg A$, such that $A \notin Cn\mathbf{X}$ and $\neg A \notin Cn\mathbf{X}$. Both sets $\mathbf{X} \cup \{A\}$ and $\mathbf{X} \cup \{\neg A\}$ are both consistent, and thereby have the Lindenbaum property. By the condition of consistency, we have two models **M**, **M′** such that $A \in$ **VER**(**M**), but $\neg A \in$ **VER**(**M′**). But **M** can be further extended to a model **extensionalM″** in which A is false. In general, expansions to the Lindenbaum sets must preserve consistency, but not truth. The identity of true (false) sentences appears to be an outcome of semantic principles used in the construction of **SDT**, because any model of a set of sentences automatically acts as the model of each of its elements. This fact results from the principle that a set (collection) of sentences is true provided every sentence in it is also true. On the other hand, we can construct restrictions (reducts) to be models of particular sentences, and this procedure is governed by the syntactical structure of considered formulas. To return to falsehoods, perhaps it is interesting that such constructions allow speaking that false sentence are such in models, although we need not to say which "parts" of models are semantic correlates of falsehoods. Thus, it is enough to say that if A is true in a model **M**, its negation is false in that model. To avoid misunderstandings, let me add that reducts and expansions in the above sense are something different than enlargements and restrictions of models considered in model theory (see por. Chang, Keisler 1990, pp. 21–22). Generally speaking, the latter increase or decrease the universes without changing **L**, although they cut or expands alphabets. It is important to note that the topics discussed in this digression are irrelevant for the construction of **SDT**. See also Sect. 8.7 for another attempt to construct denotations of sentences.►

8.3 Truth and Arithmetic or Truth Arithmetized

8.3.1 Arithmetic

Leopold Kronecker famously said (in 1886), "God made the integers, all else is the work of man". This statement can be considered as a tribute to the glorious role of the theory of natural numbers in mathematics. The great success of the arithmetization of analysis influenced essentially the foundations of mathematics. In fact, all foundational currents build on this distinguished position of integers in the entire mathematics. Logicism intended to reduce arithmetic to logic, intutionism considered natural numbers as the basic mathematical reality existing in the human mind, but formalism defined finite methods as naturally selected from the stock of arithmetical procedures. Before 1931, nobody expected that Peano arithmetic—one of the patterns of perfectly axiomatized mathematics—would produce deep problems afforded provided by the so-called limitative theorems. In fact, these results shocked the mathematical community, and are of the utmost importance for the theory of truth. The presentation of arithmetic below is very sketchy (see Hájek, Pudlák 1998 for an extensive treatment of this theory). In particular, I omit, except for general remarks (following Chap. 5, Sect. 5.3), details related to recursivity. It is enough to understand this concept as computability in the intuitive sense (in fact, it is proposed in (**ChT**) (the Church thesis; see Chap. 5, Sect. 5.3), that is, as computable or decidable step by step (algorithmically). The concept of recursivity is syntactic, not semantic. The main problem on the borderline of recursivity and semantics is the question of whether semantic regularities are recursive or not. Fortunately, the answer does not require entering into the formal machinery of recursion theory.

The first-order arithmetic of natural numbers, **AR** for brevity, plays the principal role in this section (I follow Murawski 1999, Murawski 1999a, pp. 97–103, but I use a different notation).

(**DG6**) The name 'Peano arithmetic' is used as equivalent to **AR**. However, the original version of Peano arithmetic is second-order. See Chap. 5, Sect. 5.4 for remarks about first and second order formalizations.►

AR is extension of **FOL.** The alphabet of **FOL** is supplemented by adding: one individual constant 0 (zero), one unary function symbol *seq* (being a successor of), and the two binary function symbols + (addition) and · (multiplication). Assuming **FOL** as the logical base of **AR**, the latter theory has the following specific axioms:

(**AR1**) $\forall x(seq(x) = seq(y) \Rightarrow x = y))$;
(**AR2**) $\forall x \neg(0 = seq(x))$;
(**AR3**) $\forall x(x + 0 = x)$;
(**AR4**) $\forall x(x + seq(y) = seq(x + y))$;
(**AR5**) $\forall x(x \cdot 0 = 0)$;
(**AR6**) $\forall x(x \cdot seq(y) = x \cdot y + y))$;
(**AR7**) $P(\underline{0}) \wedge \forall x(P(x) \Rightarrow P(\mathrm{seq}(x)) \Rightarrow \forall x P(x))$.

Assume the set of natural numbers **N** is to be characterized. (**AR1**)–(**AR2**) say that equal successors are successors of equal natural numbers and that **0** is not a successor of any natural number. Both these axioms imply that $\mathbf{0} \in \mathbf{N}$ (note that bold characters refer to numbers, not to numerals; **seq** refers to the function denoted by *seq*). The next four axioms define + and · as operations defined over **N**. The last axiom provides the scheme of mathematical induction, which states that if a property *P* can be predicated about **0**, and if *P* holds for $\mathbf{n_k} \in \mathbf{N}$, then it holds for **seq** ($\mathbf{n_k}$), and it can be attributed to every natural number. We think of $\mathbf{n_0}$ (= **0**) $\in$ **N**, $\mathbf{n_1}$ (= **1**) $\in$ **N**, $\mathbf{n_2}$ (= **2**) $\in$ **N**,… as just the natural numbers.

Now we have that **AR** = *Cn*{(**AR1**)–(**AR7**)}; **ARQ** = *Cn*{(**A1**)–(**AR6**)} (system **Q**, the Robinson arithmetic), **AR** without the induction axiom); **ARP** = *Cn*{(**AR1**)–(**AR4**), (**AR7**)} (Presburger arithmetic, **AR** without multiplication),; **ARS** = *Cn*

{(AR1–(AR5), (AR6, AR7)} (Skolem arithmetic, **AR** without addition). If we take intutionistic logic as the logical basis (it consists in replacing the axiom $\neg\neg A \Leftrightarrow A$ by $A \Rightarrow \neg\neg A$), we obtain **ARH** (Heyting arithmetic). The structure **Ω** = <**N**, **0**, **seq**, **+**, ·> is the standard natural model of **AR**. The same structure is also a model for **ARQ**. Models for **ARP** and **ARS** result in omitting **+** or · **,** respectively. The issue of the model for **ARH** is more complicated owing to specific features of intuitionistic model theory. It is enough to observe that if **ML** is governed by classical logic, but the semantics for intuitionistic logic is based on Boolean algebra with pseudo-complement (it is the algebraic counterpart of intuitionistic negation), the resulting structure is denoted by the symbol $\mathbf{\Omega}^{\mathbf{int}}$. Although this structure is not satisfactory from the intuitionistic point of view, $\mathbf{\Omega}^{\mathbf{int}}$ can function as the classically perceived model of **ARH**.

8.3.2 ω–Concepts

Let the symbol **Th** refers to a consistent theory. Recalling **(Df15)** in Chap. 5, **Th** (provided it formalizes the arithmetic of natural numbers) is ω-consistent (**Th** ∈ ωCon) relative to the sequence $t_1, t_2, t_3, \ldots, t_k, \ldots$ of its terms if and only if for any $A(x)$ (where x is free in $A(x)$ holds: $\forall t \in \{t_1, t_2, t_3, \ldots, t_k, \ldots\} A(x/t_k) \in \mathbf{X} \Rightarrow \neg\exists x \neg A(x) \in \mathbf{Th}$). In other words **Th**, is ω-consistent if and only if for any $A(x)$, if **Th** $\vdash A(0)$, **Th** $\vdash$ A(1), …, then ¬(**Th** $\vdash \exists x \neg$A(x); **Th** is ω-inconsistent in the opposite case. For any ω-consistent theory **Th**, it is not the case that **Th** proves that every **n** ∈ **N** possesses a property P, and proves that some natural number does not have this property. If the rule

$$(\omega)\ \mathbf{Th}\vdash A(0), \mathbf{Th}\vdash A(1), \ldots \vdash \forall x A(x),$$

is added to **AR**, it becomes ω-consistent. This rule (called the ω-rule) is infinitary because, generally speaking, it can have infinitely (denumerably) many premises. Now if **Th** $\vdash^{\boldsymbol{\omega}} A$ (A is deducible from **Th** via the ω-rule), the logical consequence Cn^{ω} is not compact because the validity of **Th** $\vdash^{\boldsymbol{\omega}} A$ does not require that there is a finite set **X** ⊂ **Th** such that **X** $\vdash A$. **Th** is called ω-complete if and only if for any formula $A(n)$, if $\mathbf{M}^{\mathbf{Th}} \models A(0)$, $\mathbf{M}^{\mathbf{Th}} \models A(1)$, $\mathbf{M}^{\mathbf{Th}} \models A(2)$, …, then $\mathrm{M}^{\mathbf{Th}} \models \forall x A(x)$; **Th** is ω-incomplete in the opposite case. If a theory **Th** is ω–complete, it has at least one ω-model, that is a structure in which every object is labelled by a numeral. All ω-concepts apply to theories which have numerals (or their counterparts in alphabets). The concepts of ω–consistency and ω-completeness can be generalized (see Grzegorczyk 1974, pp. 306–311) to descriptive consistency (d–consistency) and descriptive completeness (d-completeness). Both can be applied to theories with a non-denumerable number of individual constants—suitable definitions are straightforward by slight changes in the previously given definitions. A theory **Th** is constructive with respect to a sequence of terms (individual constants) if and only if for any formula $A(x)$ with one free variable the following holds

(22) **Th** $\vdash \exists xA(x) \Rightarrow$ **Th** $\vdash A(x/t)$, for some term t of **Th**.

Thus, every existential sentence of **Th** has its exemplification in the form of a sentence $A(t)$.

Relations between ω-consistency, and usual consistency, ω-inconsistency and usual consistency (as defined in Chap. 5 (**Def15d–e**) can by displayed (see Woleński 2010) by diagram (**D1**) (see Chap. 4, Sect. 4.8). We take into account the points α, β, γ, δ, ε and ζ. The interpretation is as follows: α—**Th** is ω-consistent (**ωCONS** (**Th**); β—**Th** is inconsistent, (**INCONS** (**Th**)); γ—**Th** is consistent (**CONS**(**Th**); δ—**Th** ist ω-inconsistent (**ωINCON** (**Th**); ε—**CONS**(**Th**) ∨ **INCONS** (**Th**); ζ—**CONS**(**Th**) ∧ **ωINCONS**(**Th**); moreover **Th**$^{\neg}$ (the negation of **Th**) is defined as $\exists A$(**Th** $\vdash A \Leftrightarrow$ **Th**$^{\neg} \vdash \neg A$. We have the following principles (they also hold for d-consistency and d-inconsistency when replaced for ω-cases).

(23) (a) **ωCONS**ω (**Th**) ⇒ **CONS**(**Th**);
(b) **INCONS** (**Th**) ⇒ **ωINCONS**(**Th**);
(c) ¬(**ωCONS**(**Th**) ∧ **INCONS**(**Th**));
(d) **CONS**(**Th**) ∨ **ωINCONS** (**Th**);
(e) ¬(**CONS**ω (**Th**) ⇔ ¬(**ωINCONS**(**Th**));
(f) ¬(**CONS** (**Th**) ⇔ **INCONS** (**Th**));
(g) **ωCONS** (**Th**) ⇔ **INCONS**(**Th**$^{\neg}$);
(h) **ωCONS** (**Th**) ∨ **INCONS** (**Th**) ∨ **CONS**(**Th**) ∧ **ωINCONS**(**Th**).

These facts follow from the **D1**-logic. Perhaps (1c) is surprising, because we would expect that no theory could be consistent and ω-inconsistent at the same time (see the next section for the importance of this possibility). The universal closures of the cases located at ζ and ε are false, but the former has no interesting interpretation except for the situation of the omniscient being who could effectively decide about any **Th**, whether it is ω-consistent or inconsistent. The points κ and λ are omitted, because if they are interpreted as '**Th** is true' and '**Th** is false', respectively, there is no general logical connection between truth and ω-consistency or falsehood and ω-inconsistency. This assertion does not mean that these properties of theories are completely separate, but only that they should be established in concrete cases. In general, it can happen that a theory **Th** is ω-inconsistent and consistent. This second property implies that it has a model in which its theorem are true. I will return to this question in Sect. 8.6.

Which concept—ω-consistency or consistency—formally and adequately represents our intuitions associated with consistency? The same applies to concerns inconsistency for we have a choice between ω-inconsistency and the notion of **ωINCONS** in order to formalize our intuitions. At the moment, we only know that **ωCONS** is stronger than **CONS**, similarly as **ωINCONS** is weaker than **INCONS**. Consequently, the assertions '**Th** is ω-consistent' and '**Th** is inconsistent' are contrary, but not contradictory, whereas the statements '**Th** is consistent' and '**Th** is ω-inconsistent' are complementary. If **Th** is ω-inconsistent, it contains theorems of the type $\exists xA(x)$ and $\exists x\neg A(x)$ ($\Leftrightarrow \neg\forall xA(x)$), and the problem stems from the fact that the former applies to every numeral. If **Th** does not have the ω-rule the two sentences mentioned are not contrary—and even can be true (see the next section). If the ω-rule is added, the situation changes fundamentally, because ω-consistency reduces to consistency, and ω-inconsistency to inconsistency. These remarks suggest that ω-consistency does not explain the intuitive concept of consistency and same should be said about ω-inconsistecy as a substitute for inconsistency. The ω-concepts are rather applied depending on the strength of particular models and theories formulated to describe them. If we are dealing with ω-models, as in the case of **AR**, the use of ω-concepts refines investigations on metamathematical properties (see below).

What about a connection between constructivity (in the above sense), d-completeness and d-consistency? We have the following statements (Grzegorczyk 1974, p. 310):

(24) (a) If **Th** is constructive and consistent, it is d-consistent;
(b) If **Th** is (syntactically) complete and d–consistent, it is d-complete;
(c) If **Th** is consistent, complete and d-complete, it is constructive.

Grzegorczyk suggested that if we assume that all natural numbers are named by numerals, then this set is d-complete, d-consistent and constructive with respect to the sequence of all terms naming particular objects from $\mathbf{U}^{\boldsymbol{\Omega}}$, and **SDT** correctly characterizes arithmetical truths. According to Grzegorczyk, an example is provided by describing the truths of **AR** via the use ω-consistency, constructivity and ω-completeness. However, constructivity and ω-concepts are syntactic notions—which do not fully capture the concept of truth. For instance (see next section), sets of truths in non-standard models of **AR**, can have no ω-properties. Hence, it seems that a weaker assertion is correct, namely that some truth-sets are d-consistent, d-complete and constructive in some cases, in particular, in canonical models.

8.4 Limitative Theorems

As I occasionally noted in this section, **AR** (and every theory **Th** sufficiently rich to formalize **AR**) satisfies so-called limitative theorems (see also Kotlarski 2004). They are as follows (I omit the Löwenheim–Skolem theorems and the Lindström theorem; see Chap. 5, Sect. 5.3):

(**TG1**) (The first Gödel incompleteness theorem)
If **AR** is consistent, it is syntactically incomplete, that is, there are $A, \neg A \in \mathbf{L_{AR}}$, such that $\neg(A \in Cn\mathbf{Ar})$ and $\neg(\neg A \in Cn\mathbf{X})$;

(**TG2**) (The second Gödel incompleteness theorem)
The formula **CONS**(**AR**) expressing that $\mathbf{AR} \in \mathbf{CONS}$ (arithmetic is consistent) is not provable in arithmetic, that is, '**CONS**(AR) $\notin Cn\mathbf{AR}$;

(**TT**) (The Tarski undefinability theorem)
The set $\mathbf{VER}(\boldsymbol{\Omega})$ is not definable in **AR**;

(**TC**) (The Church undecidability theorem)

AR is undecidable.

These theorems assert something mathematical about **AR** as a formal system. Their proofs require advanced mathematical devices which exceed the scope of this book. Thus, I limit myself to some basic remarks [at the moment I neglect (**TT**)].

(**TC**) follows from the assertion that **FOL** is undecidable [see Chap. 5(25)] and says that no algorithmic method is available for deciding whether an arbitrary A, $A \in \mathbf{L_{AR}}$ is a theorem of **AR**. (**TG1**) is interesting from the point of view of **STT**. Consider the sentence

(*) This sentence is unprovable.

Assume that (*) is true. If (*) is such, it is unprovable. But if (*) is false, it is also unprovable, because logic does not prove falsehood. Consequently, (*) is unprovable. This reasoning is informal. Gödel invented the method of arithmetization of syntax. $\mathbf{L_{AR}}$ is the language of formalized **AR**. $\mathbf{ML_{AR}}$ is an informal (mathematical) metalanguage in which we can formulate such assertions as '**AR** is consistent', '**AR** is complete', etc. The arithmetization consists in a coordination of expressions of $\mathbf{ML_{AR}}$ with Gödel numbers defined in **AR,** and transforms expressions of the former into linguistic items of the latter. Consequently, metamathematical statements about **AR** become sentences of **AR** and have arithmetical proofs, if any. Denote arithmetical version of (*) by the letter G. We obtain that if **AR** is consistent, then $\neg(\mathbf{AR} \vdash G)$ and if **AR** is ω-consistent, then $\neg(\mathbf{AR} \vdash \neg G))$. If G is made more complicated, the given version of (**TG1**) is provable (I will assume that). As far as (**TG2**) is concerned, we first show that the formulas **CONS** (**AR**) and ¬**CONS**(**AR**) can be taken as G and $\neg G$, respectively. Since we assume that **AR** is consistent, we have that $\neg(\mathbf{AR} \vdash \mathbf{CONS}(\mathbf{AR}))$, that is that **AR** does not prove its own consistency. All reported results hold for **ARQ** and **ARH**. **ARP** and **ARS** are decidable and complete. My further remark concerns **AR** (assuming that it is consistent) and, if it so indicated—**ARH**.

A couple of supplementary remarks are in order. The sentence (*) is self-referential, but—not paradoxical. Secondly, replacing a stronger condition (ω-consistency in this case) by a weaker one (consistency) makes the proof more complicated. Thirdly, let G and $\neg G$ be unprovable arithmetical sentences. Since we assume that **AR** is consistent, both new systems, namely $\mathbf{AR} \cup \{G\}$ and $\mathbf{AR} \cup \{\neg G\}$, are also consistent. In other words, we can regard these systems as different extensions of **AR** with new axioms, namely G and $\neg G$, respectively. A similar construction can be performed for $\mathbf{AR} \cup \{G\}$ and $\mathbf{AR} \cup \{\neg G\}$, and new unprovable sentences appear. Since we can repeat this procedure ad infinitum, the set of theorems of **AR** is not finitely axiomatizable [but it is recursively axiomatizable). Since the intuitive proof of (**TG1**) is semantic, its generalization entails (assuming (**BI**)] that there exist true, but unprovable arithmetical sentences [this formulation expresses the semantic version of (**TG1**)]. Adding the ω-rule results in the semantic completeness of **AR** (relative to **Ω**), but it still remains syntactically incomplete. Semantic completeness can be also achieved by taking all theorems of **AR** (without axiomatization), but this is a rather artificial manoeuvre. Since arithmetization converts the semantic proof of (**TG1**) into a syntactic one, one can say that this procedure provides a representation of semantics in syntax. The problem that arises is of how adequate this representation is. To anticipate, due to (**TT**), the representation of semantics in syntax is only partial. In other words, the set of provable arithmetical sentences is a proper subset of the truths of **AR**.

(**DG7**) Gödelian sentences are also called 'undecidable'. One should sharply distinguish the expressions 'the set **X** is undecidable' [like in (**TC**)] and 'a sentence *A* is undecidable relative to the set **X**'. Whereas the former means that no algorithm decides whether a formula belongs to **X** or not, the latter points out that a given sentence *A* is not provable from **X**'. I will use 'unprovable', but 'undecidable' will appear in many quotations.►

(**DG8**) The adjectives 'constructive' or 'effective' have various meanings in the context foundational discussions. Hence, some remarks are in order here. First of all, Grzegorczyk's (see above) use is much too specific and should be omitted. Secondly, since I assume classical logic, the intuitionistic approach is too restrictive. Thirdly, Hilbert's finitism is also too restrictive. Fourthly, (**ChT**) can be rephrased as the statement that recursivity = constructivity. Unfortunately, these explanations do not generate a univocal proposal. Even if we assume that the concept of constructivity should be explicated inside **AR** via the rephrased (**ChT**), this suggestion must be made more precise. Consider various subsystems of **AR**. **ARP** and **ARS** are decidable and complete (syntactically and semantically). **ARQ** is undecidable and incomplete in the Gödel sense, but—finitary axiomatizable. It immediately entails that the axiom of induction (**AR7**) leads to not having of a finite in the case of **AR**. Hence, one might propose that **ARQ** adequately accounts for constructivity seen from the classical point of view, but the weak point of this proposal consists in the fact that this theory does not suffice for representing all recursive functions in it. This immediately shows that the equality 'recursivity = constructivity' goes beyond **APQ**. On the other hand, every recursive function can be represented in **AR**. This theory appears as the critical point as far as the issue concerns constructivity, because its extensions (e.g. via adding the ω-rule) are not constructive beyond reasonable doubts. The above data seem to suggest that we have various degrees of constructivity (see Mostowski 1955). If so, we need a convention concerning the borderline between what is constructive and what is not. It seems that **AR** appears as the minimal constructive scheme in the sense that its extensions are surely non-constructive. This convention is motivated by the situation of **STT**, because the extent to which truth can be arithmetized specifies also how this concept can be made constructive via the arithmetization of syntax. This leads to the suggestion that (first-order) syntax is constructive, but the problem of the constructivity of semantics is open (in fact, it has a negative solution).►

(**DG9**) John B. Rosser (see Rosser 1936) replaced the condtion of ω-consistency by consistency. This step was possible due to introducing a new versiion of Gödel sentences. Informally speaking, the Rosser sentence RS says that is provable via some coded proof such that there is no smaller coded proof of the negation of RS. However, this result does not chan ges the situation of sentences **CONS**(**AR**) and ¬**CONS**(**AR**), more intuitive than RS and its negation.►

8.5 The Undefinability of Truth

(**TT**) has a special importance for analyzing the concept of truth. Its informal explanation is as follows. Assume that $\mathbf{STT}^{\mathrm{L}}$ is a correct (consistent) truth-theory for **L** formulated in this language and that a formula *A* is a formula (of **L**) which says '$\mathbf{STT}^{\mathrm{L}}$ does not defines truth for **L**′. Firstly, assume that *A* is true. If $A \in \mathbf{Tr}(\mathbf{L})$, truth is undefinable by *A*. Now, *A* is not a theorem of $\mathbf{STT}^{\mathbf{L}}$, that is $\neg(\mathbf{STT}^{\mathbf{L}} \vdash A)$ (or $A \in Cn(\mathbf{STT}^{\mathbf{L}})$. This assertion is justified by the reductio argument. Assume that $\mathbf{STT}^{\mathbf{L}} \vdash A$. Hence, $(\neg A \notin Cn(\mathbf{STT}^{\mathbf{L}})$. Hence, $\neg A$ can be either false or independent of $\mathbf{STT}^{\mathbf{L}}$. The first-case is impossible, because it would mean that $\mathbf{STT}^{\mathbf{L}}$ defines truth for **L**, contrary to our assumption. Thus, the second possibility remains, namely that $\mathbf{STT}^{\mathbf{L}}$ does not define truth for **L**. Secondly, assume that *A* is false. This means that $\mathbf{STT}^{\mathbf{L}}$ defines truth of **L**. However, it is impossible, because *A* would be a false theorem of $\mathbf{STT}^{\mathbf{L}}$, but we assumed that this theory is correct. Thus, we informally proved that $\mathbf{STT}^{\mathbf{L}}$ does not define the truth- predicate for **L**.

The following form of **TT** is convenient for its proof (the symbol $\mathbf{T}(A)$ means '*A* is true'):

($\mathbf{TT}^{\mathbf{1}}$) There is no formula $\mathbf{T}(A) \in \mathbf{L}_{\mathbf{AR}}$ such that for any $A \in \mathbf{L}_{\mathbf{AR}}$, $\mathbf{AR} \vdash A \Leftrightarrow \mathbf{T}$ ('*A*').

Informally, ($\mathbf{TT}^{\mathbf{1}}$) says that no formula of the type $\mathbf{T}(A)$ codes the truth-definition for **AR**. For proof, we need the fixed-point lemma:

(**FPL**) If $A(x) \in \mathbf{L}_{\mathbf{AR}}$ and $A(x)$ has one free variable, then $\exists B \in \mathbf{L}_{\mathbf{AR}}$ ($\mathbf{AR} \vdash B \Leftrightarrow A$ ('*B*').

Proof of ($\mathbf{TT}^{\mathbf{1}}$): Assume that there is a formula mentioned in the first part of ($\mathbf{TT}^{\mathbf{1}}$). By (**FPL**), there exists a sentence formula *L* such that $\mathbf{AR} \vdash L \Leftrightarrow \neg\mathbf{T}$('*L*'). By our assumption, we obtain $\mathbf{AR} \vdash \mathbf{T}(L) \Leftrightarrow \neg\mathbf{T}(L)$, but it conflicts with consistency of **AR**.

(**DG10**) The symbol *L* that occurs in the above reasoning was chosen intentionally. The content of the equivalence $L \Leftrightarrow \neg\mathbf{T}$('*L*') allows us reproduce the Liar sentence. Consequently, if the truth-predicate for **AR** were defined in this theory, it would lead to **LP**. That is an important fact because it provides evidence that **LP** is not a linguistic curiosity, but has a definite (meta)mathematical meaning. Thus, the fact that **LP** blocks truth-definition in natural language appears as not accidental, but displays something essential. In a sense, (**FPL**) shows that self-referentiality is not mysterious (see Smorynski 1985, p. VIII) phenomenon and can be normalized by some mathematical tools (Smullyan 1994 is an extensive study about fixed-points, diagonalizations and self-reference).►

(**DG11**) Perhaps it is interesting to point out that (**FPL**) is used (see Cook 2014) in explaining the nature of the Yablo paradox (see Chap. 7, Sect. 7.6). Since, as I noted in the previous paragraph, this lemma, informally shows how to deal with self-reference without causing antinomies, this is an additional argument that the

Yablo paradox should not be considered as arising without introducing self-reference. In fact, is demonstrates that not all **T**-equivalences are provable in a theory in which self-referential sentences occur. Another important observation is that inspecting (**FPL**) shows immediately that this lemma provides a metamathematical basis for **CT**. Formally speaking, (**TT**$^{\mathbf{1}}$) entails that the predicate **T**, to be a truth-predicate for **AR** must satisfy the condition that for any $A \in \mathbf{L_{AR}}$, $\mathbf{AR} \vdash A \Leftrightarrow$ **T**('A'). Finally, the proof of (**TT**$^{\mathbf{1}}$) proceeds by *reductio ad absurdum* and as such is not constructive. In general, informal as well as formal proof of **TT** explicitly shows the role of self-referrntial constructions (or diagonalization) in metamathematics.►

The theorem (**TT**) also has other formulations (see Tarski, Mostowski, Robinson 1953, p. 40, Smullyan 1992, p. 27, p. 99, p. 104, Mendelson 1997, pp. 216–217, Murawski 1999a, Beeh 2003, Pantsar 2009; see Sect. 8.8 for historical remarks):

(**TT**$^{\mathbf{2}}$) The set of Gödel numbers of true sentence of **AR** is not arithmetical.
(**TT**$^{\mathbf{3}}$) If **Th** ($\mathbf{AR} \subseteq \mathbf{Th}$) is consistent and the diagonal function is definable in it, **Th** has no the truth-predicateTruth-predicate.

These two versions of the undefinability are mutually interconnected. To see this, assume that a unary function f (this special case can be easily generalized) is definable in **Th**, if the sentence

(*) $\forall n \exists m(\mathbf{Th} \vdash fn = m)$

holds in this theory. Furthermore, the function f is expressible in **Th**, if it is definable in it. Consider the set **X** of sentences and the mapping $\mathbf{d}: \mathbf{X} \to \mathbf{X}$ such that for any $A \in \mathbf{X}$, $A = \mathbf{d}A$ (the diagonal function). The connection of this construction with (**FPL**) is illustrated by the following diagram:

	$\mathbf{d}A_1$	$\mathbf{d}A_2$	$\mathbf{d}A_3$	...	$\mathbf{d}A_i$	.
A_1	+					
A_2		+				
A_3			+			
⋮						
A_i					+	
.						

Suppose that **T** as a truth-predicate is definable in the theory **Th**. We can think that the equality $A = \mathbf{d}A$ is a version of the scheme $A \Leftrightarrow$ **T**('A') (**T**-scheme). Definability of **T** in **Th** would mean that every concrete **T**-equivalence—that is, marked by the symbol + – just belongs to *Cn***Th** (or simply to **Th**, because it is a theory that is a set of sentences closed by the consequence operation). Provided that the diagonal function is also definable in **Th**, there exists an A such that $\mathbf{AR} \vdash \mathbf{T}(A) \Leftrightarrow \neg\mathbf{T}(`A')$. However, **LP** appears immediately in this situation. Consequently, if the diagonal function is definable in **Th**, then **T** is not. This conclusion admits a partial definition of truth [see (**DG9III**)]. Thus, in the light of (**TT^3**), **T**(x) is either not defined for all arguments or does not satisfy Tarski's condition of adequacy saying that every concrete **T**-equivalence is a derivable in **STT** (it is a concrete example **of Th**. Both considered properties—that is, definability of the diagonal function and adequacy of the truth-definition—can be reconciled in the methatheory, but not in **Th** itself. Having (**TT^3**), we can obtain (**TT^2**) by observing that, in the opposite case, the set of arithmetical truths would be arithmetically expressible. The above reasoning shows that intuitive explanations about **LP** and **T**-scheme are precisely represented in metamathematical considerations.

Since concepts of truth and models are very closely associated in metamathematics, this dependence opens yet another possibility. Thus, we have the next (and the last in this book) version of (**TT**):

(**TT^4**) If **M** is a model of **AR**, then **M** is not definable in **AR**.

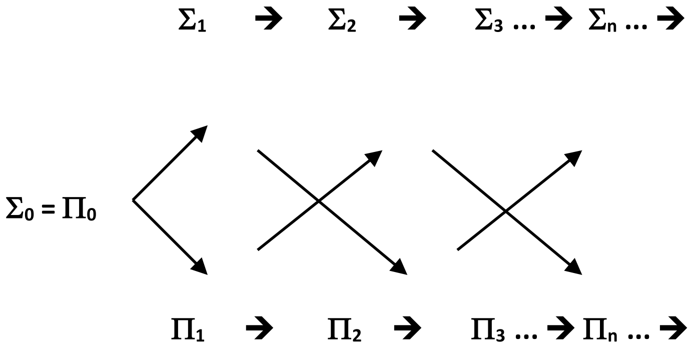
The justification of this version appeals to the fact that every consistent set of sentences has a model [see Chap. 5(26)]. If the model of **AR** were defined in this theory, it would provide a proof of '**CONS**(**AR**)', contrary to (**TG2**). This version is particularly interesting for the analysis of **SDT**, because it makes explicitly use of the concept of a model. Of course, **Ω** is the model of **AR** which is of the utmost importance, because it displays our natural understanding of natural numbers and operations performed on it. Thus, we come back to (**TT**) in this case saying that models of **AR** (including **Q** are not definable in the True Arithmetic). Still another theorem is interesting (the Askanas theorem; I follow Smullyan 1992, pp. 114–115). Assume that a formula $Q(P)$ means '$P(x_k)$ expresses the set **VER**(A)' and is arithmetized. By (**TT**), $Q(P)$ is false, so its negation $\neg Q(P)$ is true. The Askanas theorem says:

(25) $\neg Q(P)$ is provable in **AR**.

Adding $\neg Q(P)$ to **AR** does not produce inconsistency, contrary to adding **AR** $\cup$ {$Q(P)$}; (25) is similar to (**TG2**), but it immediately shows that truth is a more demanding and richer property than consistency.

(**TG1**) in its syntactic formulation and (**TT**) differ in their nature, because the latter is semantic; it cannot be expressed without using the concepts of truth and definability. If we take into account the semantic version of the former, both concern truth. Gödel's main conceptual achievement consisted in showing that the concepts of truth and proof are not equivalent, because there exist sentences true but unprovable. (**TT**) leads to the same consequence as showing that the set of provable sentences of **AR** is a subset of its truths (Tarski 1939). Formally speaking, we have the strong inclusion: Cn**AR** $\subset$ **VER**(**Ω**). In fact, Tarski's method of proving (**TT**) immediately leads to the proof of the semantic (**TG1**). On the other hand, this method does not determine the syntactic form of Gödelian sentences. This limitation is overcome by the arithmetization of syntax, which opens the way for a syntactic proof of (**TG1**). Gödel's original proof is effective (constructive in the sense adopted in this book), because it is carried out in **AR**. On the other hand, (**TT**) cannot be proved via methods accessible in **AR**. Concluding these comparative remarks, the difference between (**TG1**) and (**TT**) is essential. The former is based on construction of unprovable sentences and effectively proves that such sentences exist, but the latter only implies that they exist. Consequently, we should not say that (**TT**) is stronger than (**TG1**), but rather that Tarski invented a general—but not effective method—of proving some limitative metamathematical phenomena. This circumstance nicely display an essential feature of semantic (in the sense, of formal logical semantic) methods, namely their non-constructivity.

A precise approach to the above problems is formally available due to the so-called arithmetical hierarchy (**AH**). It considers sets of concepts defined by formulas prefixed by sequences of quantifiers. Speaking generally, we have $\mathbf{\Sigma^0}$-classes, covering concepts defined by formulas beginning with the quantifier $\exists$, and $\mathbf{\Pi^0}$-classes (their elements begin with the quantifier $\forall$); in the following remarks I omit the upper indices. **AH** starts with the classes $\mathbf{\Sigma_0} = \mathbf{\Pi_0}$, which cover all recursive relations. This is the simplest level of the hierarchy. The subsequent classes (levels) arise from former by adding quantifiers in such a way that the class $\mathbf{\Sigma_{n+1}}$ is determined by the class $\mathbf{\Pi_n}$ but the class $\mathbf{\Pi_{n+1}}$ arises from the class $\mathbf{\Sigma_n}$ (see Murawski 1999, Murawski 1999a, pp. 75–84 for rigorous definitions and examples). The number of quantifiers in a formula that defines a given concept is considered as the measure of its complexity. For instance, the notion of the limit of a sequence belongs to the class $\mathbf{\Pi_3}$, because the related definition begin with the prefix $\forall\exists\forall$, which is not subjected to any further reduction of the number of quantifiers. The objects related to $\mathbf{\Sigma_n}$ are computable (I omit a more precise characterization). The following diagram presents the relations holding between particular levels of **AH**:



The arrows mark (strong) inclusions in the group $\mathbf{\Sigma_k}$, $\mathbf{\Sigma_{k+1}}$, $\mathbf{\Pi_k}$, $\mathbf{\Pi_{k+1}}$; the schemata $\mathbf{\Sigma_k} \subset \mathbf{\Pi_{k+1}}$ and $\mathbf{\Pi_k} \subset \mathbf{\Sigma_{k+1}}$ indicate the form of inclusions between $\mathbf{\Sigma}$-classes and $\mathbf{\Pi}$-classes. The entire diagram illustrates the growth of complication of formulas stemming from adding universal and existential quantifiers.

The basic fact concerning the concept of truth via the arithmetical hierarchy expresses the theorem (Murawski 1999, Murawski 1999a, pp. 284–285):

(26) The set of Gödel numbers of true sentences in **VER**(Ω) is not a $\mathbf{\Sigma}$-class or a $\mathbf{\Pi}$-class.

That **VER**(Ω) cannot be placed at any point of **AH** provides a very good illustration of the gap between provability and truth. Although (see Wolf 2005, p. 149) for any *k*, for sentences of the form $\mathbf{\Sigma_k}$ or $\mathbf{\Pi_k}$ the truth-predicate of the same complexity can be introduced, but there is no single **T**-predicate which might reproduce the collection of all truths of arbitrarily high complexity. If we would like to use the language of recursive functions, we could say that that the set **VER**(Ω) is not recursively enumerable, contrary to the set of theorems of **AR**. Moreover, this circumstance illuminates the arithmetical undefinability of truth in the context of the strong completeness theorem (see (**Df15**(**g**) in Chap. 5). Although the set **VER(Ω)** and the set of **AR**-theorems have the same extension, the latter is recursively enumerable, but the former is not. The character of **VER(Ω)** means that this object cannot be described by any recursive procedure. Adding the ω-rule converts **AR** into a semantically complete theory, but does not suffice for defining truth by ω-consistency.

The most essential philosophical conclusion to emerge from the truth-undefinability phenomenon as displayed by (25) can be stated in the following way:

(**FC**) If **AR** $\subseteq$ **Th**, then **Th**-semantics is not expressible in **Th**-syntax.

In general, semantics is richer than syntax. In the case of first-order languages syntax is constructive, but semantics is not. Since I omit other languages, the forms sentence does not mean that the situation is as simple as described in the previous sentence. For example, the syntax of first-order languages with infinitely long expressions is not constructive, but the syntax of second-order languages is effective. Thus, the generalization in question concerns languages considered in the present book. Anyway, it says that semantic concepts and procedures are not constructive, even if it is sometimes possible to establish their extensional equivalence with their syntactic counterparts. Using infinitary tools as the ω-rule rule is indispensable for proving the parity of syntax and semantics, at least for **AR** and its over-systems. Yet such parity does not mean that semantic concepts are always definable by purely syntactic devices. In particular, the method of arithmetization does not embed semantics into syntax without a residuum belonging to the former. Truth and proof are flagship examples in this respect (see Tarski 1969, pp. 421–422):

> […] in the realm of mathematics the notion of provability is not a perfect substitute for the notion of truth. The belief that formal proof can serve as an adequate instrument for establishing truth of all mathematical statements has proven to be unfounded. […]. The notion of truth for formalized theories can now be introduced by means of a precise and materially adequate definition. It can therefore be used without any restrictions and reservations in metalogical discussions. […]. On the other hand, the notion of proof has not lost its significance either. Proof is still the only method used to ascertain the truth of sentences within any specific mathematical theory. However, we are now aware of the fact that there are sentences formulated in the language of the theory which are true but not provable, and we cannot discount the possibility that some such sentences occur among those in which we are interested and which we attempt to prove. Hence in some situations we may wish to explore the possibility of widening the set of provable sentences. […]. In

> doing so we use the notion of truth as a guide; for we do not wish to add a new axiom or a new rule of proof if we have reason to believe that the new axiom is not a true sentence, or that the new rule of proof when applied to true sentences may yield a false sentence. In the enriched theory the set of provable sentences is more comprehensive than in the original one, but it still does not contain all true sentences. This process of extending a theory may of course be repeated arbitrarily many times. The notion of a true sentence functions thus as an ideal limit which can never be reached, but which we try to approximate by widening gradually the set of provable sentences.[…]. There is no conflict between the notions of truth and proof in the development of mathematics; the two notions are not at war, but live in a peaceful coexistence.

The peaceful coexistence alluded to in the last sentence of the last quotation has a good illustration in the fact that a proof of the unprovability of a given sentence directly provides reasons for its truth or falsity (see Tarski 2001; this is a Polish translation of Tarski's lecture in Princeton in 1946; the original in Bancroft Library in Berkeley). Take, for example, the formula '**CONS**(**AR**)'. By (**TG2**) it is unprovable just as its negation is not as well, but informal explanation suggests that the former statement is true in $\mathbf{\Omega}$. Tarski's semantic method of proving the unprovability of consistency (and of any other Gödelian sentence) provides arguments for the truth of '**CONS**(**AR**)' that is fully acceptable from the mathematical point of view.

(**DG12)** According to Hintikka (see Hintikka 1991, Hintikka 1996, Hintikka 1997, Hintikka 1998) truth for **L** (first-order) is definable in the same language, assuming that we use so-called **IF**-logic. Without getting into details (see Woleński 2006a for a criticism of this approach) **IF**-logic is not compositional and implicitly uses second-order concepts. If so, it does not provide a counter-example against (**UT**). If the truth-predicated is defined in **IF**-theories, for instance, for first-order arithmetic, it means that **IF**-tools are effectively stronger than accessible in first-order languages.►

8.6 Models of AR and Truth

The structure $\mathbf{\Omega}$ is not the only model of **AR**. Since it is a first-order theory, it has various non-isomorphic (of a different cardinality) models due to (**LST**) (see Chap. 5, Sect. 5.3). However, a more interesting case stems from (**TG1**). Assuming $\mathbf{AR} \subseteq \mathbf{Th}$, we conclude that there exist the sentences G and $\neg G$, both unprovable in **Th**. Consider the theories $\mathbf{Th_1} = \mathbf{Th} \cup \{G\}$ and $\mathbf{Th_2} = \mathbf{Th} \cup \{\neg G\}$. Since we assume that **Th** is consistent, this theory has a model. If so, the theories $\mathbf{Th_1}$ and $\mathbf{Th_2}$ are consistent as well, because adding to a theory sentences which are independent of it cannot produce any inconsistency. More formally, if **Th** is consistent, $A \notin \mathbf{Th}$, then $\mathbf{Th} \cup \{A\}$ is also consistent (of course, supplementing **Th** with contradictions is excluded). Since $\mathbf{Th_1}$ and $\mathbf{Th_2}$ are consistent they have models, let say, $\mathbf{M_1}$ and $\mathbf{M_2}$. These models have to be different because one, say, $\mathbf{M_1}$ satisfies G, but $\neg G \in \mathbf{VER}(\mathbf{M_2})$. This reasoning can be infinitely iterated, because $\mathbf{Th_1}$ and $\mathbf{Th_2}$ (and their further Gödelian extensions) are also incomplete. The situation is very well-illustrated by the diagram (Murawski 1999, Murawski 1999a, p. 287):

$$\begin{array}{ccccccc} & & & \mathbf{Th} & & & \\ & & \swarrow & & \searrow & & \\ & \mathbf{Th_1} & & & & \mathbf{Th_2} & \\ \swarrow & & \searrow & & \swarrow & & \searrow \\ \mathbf{Th_{1.1}} & & \mathbf{Th_{1.2}} & & \mathbf{Th_{2.1}} & & \mathbf{Th_{2.2}} \end{array}$$

....

Let the formulas **CONS**(**Th**) and ¬**CONS**(**Th**) serve as G and $\neg G$, respectively. If we add them to **Th**, we obtain the two new theories $\mathbf{Th_1} = \mathbf{Th} \cup \{\mathbf{CONS}(\mathbf{Th})\}$ and $\mathbf{Th_2} = \mathbf{Th} \cup \{\neg\mathbf{CONS}(\mathbf{Th})\}$. Repeating these steps gives the sequences of theories $\mathrm{TH}_1 = <\mathbf{Th}, \mathbf{Th_1}, \mathbf{Th_{1.1}}, \mathbf{Th_{1.2}}, \ldots>$ and $\mathrm{TH}_2 = <\mathbf{Th}, \mathbf{Th_2}, \mathbf{Th_{2.1}}, \mathbf{Th_{2.2}}, \ldots>$, and—of models $\mathrm{MO}_1 = <\mathbf{M(Th)}, \mathbf{M(Th_1)}, \mathbf{M(Th_{1.1})}, \mathbf{M(Th_{1.2}}, \ldots>$ and $\mathrm{MO}_2 = <\mathbf{M(Th)}, \mathbf{M(Th_2)}, \mathbf{M(Th_{2.1})}, \mathbf{M(Th_{2.2})}, \ldots>$. We can assume that the branch determined by the points determined by $\mathbf{TH_1}$ develops **AR**, but that generated by $\mathbf{MO_1}$ approximates $\mathbf{\Omega}$, that is the standard model of **AR**. Other models are non-standard. It can be proved that, due to the infinitely many pairs of unprovable sentences that extend **AR**, there are uncountably many (the continuum) non-standard models of **AR**. All are mutually non-isomorphic. In particular, **AR** has countable non-standard models. This fact goes beyond the Löwenheim–Skolem property, and is actually surprising.

Analysis of the theories $\mathbf{Th_1}$ and $\mathbf{Th_2}$ proves the existence of non-standard models of **AR** by metamathematical reasoning. This argument provides no information about the structure of such models. However, it is possible to use an arithmetical argument in order to demonstrate that **AR** has models other than $\mathbf{\Omega}$ and that they are not isomorphic with the standard one. The argument proceeds as follows. We add a new constant c to $\mathbf{L_{AR}}$ such that the sentences $c \neq 0$, $c \neq 1$, $c \neq 2$, … are true. Assume **X** is an arbitrary finite subset of **AR**, but A—an arbitrary sentence of the form $c \neq k$, where k is any natural number. Since **AR** is consistent (even true), **X** has, by compactness (see Chap. 5, Sect. 5.3), the same property. The set $\mathbf{X} \cup \{A\}$ is consistent since sentence A = '$c \neq k$' is true, and adding a true sentence to a set of true sentences cannot produce an inconsistency. Let the symbol **AR′** refers to the theory arising from **AR** by adding to it all sentences of the form $c \neq k$. Since every finite subset of **AR′** is consistent, it has a model. Using compactness once again, **AR′** also has a model, namely $\mathbf{M'} = <\mathbf{N}, \mathbf{0'}, \mathbf{c}, \mathbf{seq'}, \mathbf{+'}, \cdot'>$. Intuitively, **c** is greater than every $\mathbf{n} \in \mathbf{N}$. Thus, we have standard natural numbers and non-standard ones. In particular, the number **c** is not accessible by applying the successor function to any standard natural number. The structure of **M′** is this (Kaye 1991, p. 12).

Informally, $\mathbf{M}' = \Omega + \mathbf{M}''$, where $\boldsymbol{\Omega}$ is the standard part of $\mathbf{M}'$, but $\mathbf{M}''$ is its non-standard part—collecting the non-standard natural numbers $\mathbf{w}$, $\mathbf{w}'$, … (the signs = and **+** do not have an arithmetical sense, but play an illustrative role; $\mathbf{N_{NS}}$—is the set of non-standard natural numbers.

The basic feature that differentiates the two parts of $\mathbf{M}'$ involves the type of order generated by the relation <. In the case of **N**, the order is normal (natural, etc.)—defined by the formula $x < y \Leftrightarrow \exists z \in \mathbf{N}(y = x + z)$ with **0** as the first element. The definition of <entails that only finitely many natural numbers exist between **0** and any other **n**. The situation is different in the non-standard part because there are infinitely many objects between two different non-standard numbers **w** and $\mathbf{w}'$. Thus, the ordering of the non-standard part is linear and dense, similarly as in the set of national numbers; this ordering is of the type η. If $\boldsymbol{\Omega}$ and $\mathbf{M}'$ have different order types, they cannot be isomorphic and verify different statements. Since every non-standard model is a "sum" of $\boldsymbol{\Omega}$ and a non-standard part, the standard part appears as the initial segment of every non-standard model. Due to this circumstance, the induction axiom holds in the $\mathbf{M}'$, but it is not applicable for the non-standard segment. This fact shows that ordering in the latter must be different than standard (see above).

The existence of non-standard models complicates the analysis of the concept of truth. The completeness theorem (in version (26) of Chap. 5) implies that consistency, not truth is a counterpart of the concept of model. Although every set of true propositions is consistent, the converse of this assertion is not valid. Consequently, if **Th** is consistent and it has a model, then there exists an interpretation $\mathfrak{I}$ that makes **Th** true. However, it does not need to be an interpretation that is intuitively associated with this theory. Now, consider the theory **Th = AR** ∪ {¬**CONS**(**AR**)}. Since it is consistent, it has a model, say, $\mathbf{M}^{\neg\mathbf{CONS(AR)}}$. All axioms of **AR** and their consequences are true in this model. What about the sentence ¬**CONS**(**AR**), which means '**AR** is inconsistent', even though we assumed otherwise? Thus, **Th** ⊢ ¬**CONS**(**AR**), what means that there exists a Gödel number g of a provable formula expressing that **AR** is inconsistent—implying, for instance, the sentence 0 = 1. However, g cannot refer to any standard natural number, for if it did, this would mean that inconsistency is provable in **AR**, contrary to the initial assumption.

The mentioned supposition of ω-consistency related to the formula ¬**CONS** (**AR**) as an example of a formally unprovable sentence suffices to considering this assertion (that is, of inconsistency of arithmetic) as true in $\mathbf{M}^{\neg\mathbf{CONS(AR)}}$ (Tarski 1933a is probably the first treatment of truth, still informal, of sentences belonging to consistent and ω-iconsistent theories). Although, ω-consistency as a syntactic notion, does not provide the exact counterpart of the concept of soundness as a semantic category, but at least in the context of (**TG1**), it does block deriving falsehoods from truths. If the sentence ¬**CONS**(**AR**) is provable in our **Th** and true in $\mathbf{M}^{\neg\mathbf{CONS(AR)}}$, which is non-standard, it must be false in **Ω**. In fact, its literal meaning is difficult to express in language suited to the properties of the standard model. A possible compromise consists in saying that the considered **Th** is consistent as having a model, but, if regarded from the standard point of view, it is also ω-inconsistent (see (23) and Lindström 2002, pp. 46–47 on the role of sentences like '… is not ω-inconsistent'). This solution does not exclude that a given **Th** contains standard falsehoods. The same reasoning applies to the Rosser sentence [see (**DG9**)]. Denote it by RS. Since it is independent of **AR**, **AR** ∪ {RS} and **AR** ∪ ¬{RS} are consistent and have models. If we say that RS is true in a standard model, its negation is true in a non-standard one.

If we say that every consistent theory has an interpretation that makes it true in a model, d-consistency and constructivity (in Grzegorczyk's sense) are too strong constraints to characterize an arbitrary set of truths. The set $\mathbf{VER}(\mathbf{M}^{\neg\mathbf{CONS(AR)}})$ provides an example. Since it is consistent and ω-inconsistent, it cannot be constructive for constructivity and consistency imply ω-consistency. This suggests that consistency and completeness are minimal syntactic conditions characteristic for arbitrary sets of truths, but stronger constraints depend heavily on concrete circumstances. For instance, the set **VER**(Ω) is constructive, ω-consistent and ω-complete, but the set $\mathbf{VER}^{\neg\mathbf{CONS(AR)}}$ does not possess the last two properties. In general, **SDT** does not force truths (true sentences) in its understanding to be standard, because its correctness is associated with the **CT**, and not with deciding about the truth of particular sentences and its sources. These considerations do quite well to point out that the role of **ML** to transmit intuitions and decode the meanings of expressions. Consequently, if a theory **Th** possesses non-standard models, they cannot be identified in $\mathbf{L}_{\mathbf{Th}}$ without recurring to constraints assumed in **ML**, for instance, related to various interpretations of the object language. On the other hand, non-standard models have an essential (meta)mathematical role (see Kossak 2004), for it is possible to prove (**TT**) assuming that such structures exist. It means that the undefinability of truth is deeply associated with properties of **AR** and its models.

Now, **AR** is considered as the True Arithmetic. Why? The working mathematician has a simple answer—because t it is the set of logical consequences of (**AR1**)–(**AR7**). Since these axioms are evidently true, there is to it that's all. A more theoretical (philosophical, logical) standpoint consists in suggesting that the True Arithmetic = **VER**(**Ω**)—because it overcomes limitations imposed by the incompleteness phenomena. However, the question arises as to why other models than **Ω** are omitted. They have not only a metamathematical but also a mathematical relevance, for example, in non-standard analysis (precisely speaking what is involved are non-standard models of the reals). Perhaps one would say that there is not the first case of using rival mathematical theories (*vide*, geometry) do various problems, and **AR** and **Ω** are in a sense basic. Well, but this answer opens a new question, namely 'What does 'basic' means?" The answer that **Ω** is standard cannot be regarded as satisfactory owing to its circularity. Sometimes the adjective 'intended' appears in this context, but it requires an explanation, even if we ignore its psychological flavour. The best option would be give a formal account of being a standard model. The simplest proposal (see Gaifman 2003 for a survey) suggests that **N** is well-ordered, due (**AR7**), that is the axiom of induction. On the other hand (see Tarski 1933a) non-standard models of **AR** can be "standardized" by using infinite induction. Moreover, there is the question of whether this definition can be extended beyond **AR**; for instance, which model of set theory is standard, that with the continuum hypothesis or the one with its denial; both are defended as intuitively plausible. I will not discuss this issue, but even very simple examples show that mathematicians are inclined to consider standard cases as the simplest, the most evident, etc. I return do this question in Chap. 9, Sect. 9.3 (the problem of analyticity of **T**-sentences) and **DG8IX** on the occasion of the intelligibility of meaning.

At a certain point it becomes clear that syntactic criteria do not suffice for being standard, because even the strongest conditions of this kind provide no adequate specification of sets of true sentences—except in rather simple cases. This confirms the assertion (**FC**)—perhaps the most important conclusion of the considerations in this section—that semantics is not reducible to syntax. The above analysis of standard and non-standard models, and the fact that **SDT** does not favour the former, confirms that information stemming from **ML** for the meanings of expressions in **L** is necessary for fixing what is standard and what is not. In other words, the standard model is always such relative to an assumed (even, tacitly) interpretation of **L.** This conclusion is in complete accord with Tarski's view (more than once mentioned) that the problem of truth is meaningless for entirely formal languages. It is a very deep reason why we should speak about truth in a model, but about truth *simpliciter.* Clearly, the reference to a model is redundant in some circumstances, for instance, in the case of **AR**. It is quite possible that the distinctive role of **Ω** has its source in our psycho-physical nature, that our cognitive representation of the reality is discrete and thus conditioned calculation with natural numbers as fundamental. If so **Ω** has a special position in the epistemological perspective, but at this point we enter philosophy.

8.7 Models Constructed on Terms

Models constructed on terms are related to the following theorem (Grzegorczyk 1974, p. 307; I use the symbolism from Chap. 5, Sect. 5.2.4, but omit the function symbols and simplify indexing):

(27) If **Th** is a theory with the set of $\mathbf{CON}^{\mathbf{Th}}$ and the predicates $P_1, \ldots, P_k, \ldots$, and if **Th**
Is **Th** is d-complete and d-consistent with respect to the sequence <**c**> of all terms from $\mathbf{CON}^{\mathbf{Th}}$, then $\mathbf{Th} = \mathbf{VER}(\mathbf{M}^{\mathbf{H}})$ such that $\mathbf{M}^{\mathbf{H}} = \langle \mathbf{U}^{\mathbf{H}}, \langle \mathbf{c} \rangle, \ldots, \mathrm{P}_1, \ldots, \mathbf{P}_{\mathbf{n}} \rangle$, where $\mathbf{U}^{\mathbf{H}}$ consists of all terms from <**c**>.

$\mathbf{M}^{\mathbf{H}}$ is a model (called the Henkin model; see Henkin 1949) constructed on terms. It replaces an initial model **M**, where **Th** = **VER**(**M**). Now <**c**> is sufficiently large to provide names for every **a** ∈ **U**. Intuitively speaking, the construction of $\mathbf{M}^{\mathbf{H}}$ consists in adding as many individual constants to the set $\mathbf{CON}^{\mathbf{Th}}$ as is necessary to name all object from the initial universe **U**. Consequently, all terms in the universe $\mathbf{U}^{\mathbf{H}}$ act as their self-referential names; I will use the sign $\ulcorner t_n \urcorner$ for such in-quoting. This means that the symbol $\ulcorner t_n \urcorner$ is the name of the term t_n. For simplicity, I limit further remarks in this section to monadic predicates. We have if $A = P(t)$, then 'A' ∈ **Th** ⇔ **P**($\ulcorner t \urcorner$), and if **v**($\ulcorner t_n \urcorner$) = t_n, then 'A'∈ **VER**($\mathbf{M}^{\mathbf{H}}$) ⇔ $\mathbf{P}_k$($\ulcorner t_n \urcorner$) ⇔ 'A' ∈ **Th**. The formula 'A' ∈ **VER**($\mathbf{M}^{\mathbf{H}}$) ⇔ $\mathbf{P}_k$($\ulcorner t \urcorner$) also means that the sentence A is true in **M**, because the theories **VER**(**M**) and **Th** are d-complete and d-consistent with the respect to the same sequence of individual constants, and they have the same atomic formulas (the generalization for predicates of arbitrary arity is straightforward).

Every consistent theory **Th** has the described extension (the Henkin extension), which is maximally consistent. This is a simple consequence of This circumstance guarantees that any model **M** can be replaced by its counterpart constructed on terms, provided that the new set of terms is sufficiently extensive and that the condition of consistency is preserved. However, this initial model and **VER**(**M**) serve as starting points in the entire construction. The substitute for **M,** constructed on terms appears later as, so to speak, a by-product. This means that relations holding in **M**, under an assumed interpretation, can be reproduced in $\mathbf{M}^{\mathbf{H}}$ consists of linguistic material. The cost of this step amounts to adding a number of individual constants sufficient to obtain a d-complete theory. These terms become their own names does not change the interpretation earlier established.

The outlined construction leads to a proposal concerning denotations of sentences. As a first step, we define the Lindenbaum algebra of sentences. It divides sentences into abstraction classes (**LABC**—Lindenbaum abstraction classes, Lindenbaum sets) such that sentences A and B belong to the same **LABC** if and only if $A \dashv\ \vdash B$ (A, B are logically equivalent; the sign $\dashv$ in the context $\dashv\ \vdash$ does not mean 'is refuted'). Furthermore, $B \in |A|$ if and only if $A \dashv\ \vdash B$, $B \in -|A|$ if and only $B \in |\neg A|$ if and only if $A \dashv\ \vdash \neg B$. Further, $\neg A \vee B$, $\neg(A \wedge \neg B) \in |A \Rightarrow B|$, $\neg A\ \in$—$|A|$, etc., as well as $|A| \cup -|A| = |A \vee \neg A|$ and $|A| \cap -|A| = |A \wedge \neg A|$. These principles can be generalized for a theory **Th**. Thus, $B \in |A|^{\mathbf{Th}}$ if and only if $A\ {}^{\mathbf{Th}}\dashv\ \vdash^{\mathbf{Th}} B$. Observe that $A\ {}^{\mathbf{Th}}\dashv\ \vdash^{\mathbf{Th}} B$ is weaker than $A \dashv\ \vdash B$, because logical bi-entailment implies theoretical bi-entailment, but not conversely. The sources of theoretical bi-entailment can be various, for instance, specific axioms, meaning postulates, analyticity, etc. I will return to this issue after presenting the formal construction.

Assume that **Th** satisfies the conditions listed in (27). Replacing the term t in the formula $P(t)$ by $\ulcorner t \urcorner$ yields the expression $P(\ulcorner t \urcorner)$. The formula '$P(t)$' ∈ **VER** ($\mathbf{M}^{\mathbf{Th}}$) ⇔ **P**($\ulcorner t \urcorner$) becomes the equivalence '$P(t)$' ∈ **VER**(**M**) ⇔ $\ulcorner P(t \urcorner$), valid in $\mathbf{M}^{\mathbf{HTh}}$ In general, 'A' ∈ **VER**($\mathbf{M}^{\mathbf{Th}}$) iff $\ulcorner A \urcorner$, because $\mathbf{M}^{\mathbf{HTh}}$ is now a Henkin model for **Th**, and $A \in$ **Th**). These assignments motivate

(28) (a) If $A = \ulcorner Pt \urcorner$, then '$A$' $\in$ **VER**($\mathbf{M^{Th}}$) $\Leftrightarrow \mathfrak{J}$('$A$', $\mathbf{M^{HTh}}$) = $|P(t)|^{\mathbf{Th}}$;
(b) If $A = \ulcorner \neg B \urcorner$, then '$A$' $\in$ **VER**($\mathbf{M^{Th}}$) $\Leftrightarrow \mathfrak{J}$('$A$', $\mathbf{M^{HTh}}$) = $-|B|^{\mathbf{Th}}$;
(c) If $A = \ulcorner B \wedge C \urcorner$, then '$A$' $\in$ **VER**($\mathbf{M^{Th}}$) $\Leftrightarrow \mathfrak{J}$('$A$', $\mathbf{M^{HTh}}$) = $|B \wedge C|^{\mathbf{Th}}$;
(d) If $A = \ulcorner B \vee C \urcorner$, then '$A$' $\in$ **VER**($\mathbf{M^{Th}}$) $\Leftrightarrow \mathfrak{J}$('$A$', $\mathbf{M^{HTh}}$) = $|B \vee C|^{\mathbf{Th}}$;
(e) If $A = \ulcorner B \Rightarrow C \urcorner$, then '$A$' $\in$ **VER**($\mathbf{M^{Th}}$) $\Leftrightarrow \mathfrak{J}$('$A$',$\mathbf{M^{HTh}}$) = $|B \Rightarrow C|^{\mathbf{Th}}$;
(f) If $A = \ulcorner B \Leftrightarrow C \urcorner$, then '$A$' $\in$ **VER**($\mathbf{M^{Th}}$) $\Leftrightarrow \mathfrak{J}$('$A$', $\mathbf{M^{HTh}}$) = $|B \Leftrightarrow C|^{\mathbf{Th}}$;
(g) If $A = \ulcorner \exists x Px \urcorner$, '$A$' $\in$ **VER**($\mathbf{M^{Th}}$) $\Leftrightarrow \mathfrak{J}$($\ulcorner A \urcorner$, $\mathbf{M^{HTh}}$) = $|P(t)|^{\mathbf{Th}}$, if $A \vdash P(t)$, for some t.
(h) If $A = \ulcorner \forall x B \urcorner$, then '$A$' $\in$ **VER**($\mathbf{M^{Th}}$) $\Leftrightarrow \mathfrak{J}$('$A$', $\mathbf{M^{HTh}}$) = $\left| \ulcorner P(t) \urcorner \right|^{\mathbf{Th}}$, if for any t, $A \vdash P(t)$;
(i) 'A' $\in$ **FLS**($\mathbf{M^{Th}}$) $\Leftrightarrow$ 'A' $\notin$ **VER**($\mathbf{M^{Th}}$);
(j) 'A' $\in$ **TAUT** $\Leftrightarrow \mathfrak{J}$ ('A') = $||A| \cup -|A||$;
(k) '$\neg A$' $\in$ **TAUT** $\Leftrightarrow \mathfrak{J}$ ('A') = $||A| \cap -|A||$.

By the similarity of the initial model $\mathbf{M^{Th}}$ and the Henkin model $\mathbf{M^{HTh}}$ we can identify the denotation of a sentence A with $|A|^{\mathbf{Th}}$, that is with its denotation considered as a suitable **LABC**. The points (29b–h) appeal to definitions of logical constants in **PC** and **FOL**. Thus, two sentences which belong to the same Lindenbaum set—that is, they remain in the relation of theoretical bi-entailment—have the same denotation. Note that replacing bi-entailment by material equivalence would force the statement that two arbitrary true sentences have the same denotation. Thus, denotations of true sentences are distinguishable by intensions. The point (28i) is essential because it allows the thesis that false sentences are such in models. Definition (28) weakens somehow this consequence because it works with **LABC**. However, the intensional aspect of the entire construction does not disappear, because the initial model and its interpretation in **ML** generate **LABC.** (28 g) and (28 h) show that d-consistency, d-completeness and constructivity are important for defining denotations and truth. This circumstance can be taken as a certain justification of Grzegorczyk's view (see Sect. 8.3.2 and a qualification mentioned there) that these properties are marks of truth in models. It is not necessary to make reference to theories in (28j) and (28k) because tautologies are true in every model. Since **LABC** are not elements of models, (28) does not define denotations of sentences in models. Yet we want to say that truths refer to something in a model. However, we can add **LABC** to Henkin models (it changes nothing from the purely metalogical point of view; this step serves some philosophical aims). In other words, $\mathbf{M^{HTh}}$ can be extended to $<\mathbf{M^{HTh}}, <|\ldots|^{\mathbf{Th}} \gg$, where $<|\ldots|^{\mathbf{Th}}>$ is a sequence of **LABC** sufficiently large to cover all theoretical bi-entailments in order to provide all denotations for truths of a given **Th**. This manoeuvre gives full justification to (28) for defining truth exclusively by the content of models. Since passing from $\mathbf{M^{HTh}}$ to $\mathbf{M^{Th}}$ is always admissible, we can say that (28) suggests a metalinguistic approach to denotations of sentences in $\mathbf{M^{Th}}$ via $\mathbf{M^{HTh}}$.

Nevertheless, the above approach does not seem satisfactory from the philosophical point of view. Most philosophers, I assume, would like to say that if A is true in **M**, the model in question should be equipped with something which marks that A is true. Schematically speaking, we expect that

(29) 'A' $\in$ **VER**(**M**) $\Leftrightarrow$ **A** occurs in **M**,

where **A** symbolizes the denotation of A, due to which the sentence A obtains in **M.** Think about such items as states of affairs or situations (I do not intend to develop here a situation semantics in the sense of Barwise, Perry 1983). For instance, the object For instance, if $A = P(a)$, then the object **Pa** is the state of affairs relative to $P(a)$. Consequently, '$P(a)$' is true in **M** if and only if **Pa** is the denotation of 'A' in **M**. And we should define **A's** directly, not by the mediation of linguistic resources.

Although I appreciate philosophical claims requiring a direct insight into things, I believe that we cannot abstract from language-matters. When looking for items suitable to be denotations of sentences, we have to take a given **L** and its interpretation. Hence, I assume that **LABC'** somehow represent states of affairs. Thus, if $A = P(a)$, then the object $\Vert \mathbf{Pa} \Vert$ is a state of affairs relative to $P(a)$; this symbolism intends to indicate the connection between $\Vert \mathbf{A} \Vert$ and $|A|$. Moreover, if $A \dashv\vdash B$, then $\mathbf{B} \in \Vert \mathbf{A} \Vert$, but $\mathrm{B} \in \Vert \mathbf{A} \Vert$ if and only if $A \Leftrightarrow B$, is not admissible. Thus, **LABC** for sentences (**LABBC**$^{\mathbf{s}}$) become **LABC** for states of affairs—**LABC**$^{\mathbf{sa}}$). To imitate the extension of Henkin models by **LABC**, we add to a given **M** a sequence of **LABC**$^{\mathbf{sa}}$ in order to obtain the structure <**M**, <**LABC**$^{\mathbf{sa}}$>>. Now we can rewrite (28) with respect to a theory **Th** (I omit indices for simplicity):

(30) (a) If $A = P(t)$, then 'A' $\in$ **VER**(**M**) $\Leftrightarrow \mathfrak{I}$('$A$', **M**) = $\Vert \mathbf{Pa} \Vert$;
(b) If $A = \neg B$, then 'B' $\in$ **VER**(**M**) $\Leftrightarrow \neg$('A' $\in$ **VER**(**M**));
(c) If $A = B \wedge C$, then 'A' $\in$ **VER**(**M**) $\Leftrightarrow \mathfrak{I}$ ('A', **M**) = **inf** $\Vert \mathbf{B} \Vert \cap \Vert \mathbf{C} \Vert$;
(d) If $A = B \vee C$, then 'A' $\in$ **VER**(**M**) $\Leftrightarrow \mathfrak{I}$ ('A', **M**) = **sup** $\Vert \mathbf{B} \Vert \cup \Vert \mathbf{C} \Vert$;
(e) If $A = B \Rightarrow C$, then 'A' $\in$ **VER**(**M**) $\Leftrightarrow \mathfrak{I}$ ('A', **M**) = $\mathfrak{I}$ $(\neg B \vee C))$
(f) If $A = B \Leftrightarrow C$, then 'A' $\in$ **VER**(**M**) $\Leftrightarrow \mathfrak{I}$ ('A', **M**) = $\mathfrak{I}$ $((B \Rightarrow C^{)} \wedge (C \Rightarrow B))$;
(g) If $A = \exists xPx$, 'A' $\in$ **VER**(**M**) $\Leftrightarrow \mathfrak{I}$('$A$', **M**) = $\Vert \mathbf{Pa} \Vert$, if $\exists xP(x) \vdash P(t)$;
(h) $A = \forall xB$, then 'A' $\in$ **VER**(**M'**) $\Leftrightarrow \mathfrak{I}$('$A$', **M**) = $\mathfrak{I}$ $(\neg\exists x\neg P(\mathrm{x})$;
(i) 'A' $\in$ **FLS**(**M**) $\Leftrightarrow$ 'A' $\notin$ **VER**(**M**);
(j) 'A' $\in$ **TAUT** $\Leftrightarrow \mathfrak{I}$('$A$') = $\Vert \mathbf{A} \Vert \cup - \Vert \mathbf{A} \Vert$;
(k) 'A' $\in$ **TAUT** $\Leftrightarrow \mathfrak{I}$('$A$') = $\Vert \mathbf{A} \Vert \cap - \Vert \mathbf{A} \Vert$.

Since the Lindenbaum algebra is a lattice, (30c) and (30d) use the concepts of infinum and supremum. If we add $\mathfrak{I}$, it is not correct to say that $\mathfrak{I}$('A', **M**) = $\Vert \mathbf{A} \Vert$, but the proper rendering is that $\mathfrak{I}$('A', **M**) = $\Vert \mathbf{A} \Vert$ or $\mathfrak{I}$('A', **M**) = $- \Vert \mathbf{A} \Vert$; this reproduces (**BI**). As a result, we have only two sets of sentences—truths and falsehoods—according to the principle that equivalences are true provided that both their constituents have the same logical value. This example shows how **LABC**$^{\mathbf{s}}$ and **LABC**$^{\mathbf{sa}}$ are related to the intensional aspect of interpretations in models, although they are extensional constructions. It seems that (28) and (30) provide what is (almost, to be careful) possible to represent intensionality in the conceptual apparatus consistent with classical logic. Once again, the general explanation of bi-entailment is partially intensional due to its appealing to synonymy of expressions, for instance. Stronger results could be achieved via intensional isomorphism (see Carnap 1947) but, to say it once again, the full theory of intensionality is still an open issue. However, related conditions point out a parallelism between two approaches—linguistic (models on terms) and ontological (the algebra of states affairs). Perhaps the view (Goldstern, Judah 1995, s. XII)), that the reality is the Henkin Model goes too far, but it has some merit because if we want to speak about individuals as such, we have to name them. Logic cannot decide how many things there are.

(DG13) The approach proposed in this section will not be used in my further considerations as important for understanding **STT**, even as a correspondence theory of truth. I outlined a formal account of denotations of sentences to show that logic provides tools for an analysis such concepts as that of state of affairs.►

(**DG14**) The most general characterization of model theory is as follows (Chang, Keisler 1990, p. 1):

> Model theory […] deals with the relation between a formal language and its interpretations, or models. […]; our own usage is explained by the equation [:] universal algebra + logic = model theory.

According to this explanation, model theory is a genuine part of mathematics, and has several important applications, for instance, in non-standard analysis and algebra. In this framework, (**SDfVER**) (see Chap. 5, Sect. 5.2.4), that is, a simplified truth-definition (it does not proceed by the concept of satisfaction) is enough for investigating relations between formal (or formalized) languages and models. Functioning this device in metamathematics dos not require to say that it is a definition of truth or that it belongs to semantics, although the latter is suggested by the quoted passage from Chang, Keisler 1990. (**Df4**) points out items from models responsible for being true. Roughly speaking, whereas (**SDfVER**) explicates the content of **T**-scheme by translating its right part into set-theory, (**Df4**) offers the reason why this translation is correct. This ingredient in (**Df4**) allows presenting several intuitions (also philosophical) behind formalities—Tarski did the same, but certainly not so much. This extensive treatment of **SDT** is justified by the aim of the present book consisting in connecting logic and philosophy (see also my final conclusion).►

8.8 Historical Appendix: Gödel and Tarski on Limitative Theorems and Truth

This section (based on Woleński 2004, but with some amendments) describes various data important for the history of discovering the undefinability of truth up to 1939 (I use the abbreviation (**UT**) [not (**TT**)] for the discussion concerning the problem of priority).

8.8.1 Chronology (up to 1939; See also DG1VII)

1929

Gödel completes his Ph.D. (Gödel 1930; the original text is published in Gödel 1986; I will refer to it as to Gödel 1929) on the completeness of first-order logic. There is a difference between Gödel 1929 and Gödel 1930. The former (p. 63–69) contains various informal explanations on validity (*allgemeine Gültigkeit*) and satisfaction (*Erfüllbarkeit*). Gödel speaks (p. 65) about satisfaction (or, sometimes, realization, *Realisierung*) of logical expressions by a system of relations, when 'the sentence obtained through substitution is true". He then adds (p. 69) that satisfaction concerns "the domain in question", and finally says (p. 69): "what is to be understood by *consistent*, *satisfaction* (*realization*), *satisfiable*, and so on, is immediately clear." The last passage strongly suggests that he took these concepts as intuitively obvious. These explanations were deleted in Gödel 1930. Of course, Gödel uses the mentioned words but without comments. Some commentators (see Dreben, van Heijenoort 1986, p. 50) remark that it was Hans Hahn, Gödel's supervisor, who recommended deletions in order to do not provoke philosophical deliberations around purely logical issues.

1930

Gödel announces his (first) incompleteness theorem in Königsberg (see Gödel 1931a), no mention of (**UT**);
Gödel 1930a is published, no mention of (**UT**);
Tarski visits Vienna (see Menger 1994, Ch. XII about this visit), delivers three lectures, has discussions with Carnap (i.e., about the language/metalanguage distinction), and has a meeting with Gödel (see Feferman 1999 about the meetings of Tarski and Gödel);
Tarski announces his truth-definition in Lvov.

1931

Tarski 1930–1931 is published, no mention of (**UT**);
Gödel's letter to Tarski (see Gödel 2003a, p. 267); Gödel promises to send his Ph. D. thesis and forthcoming Gödel 1931;

Gödel's letter to Bernays (April 2); Gödel explains some points of his results;

Gödel 1931 is published. Gödel uses the concept of truth in an informal presentation of the first incompleteness theorem. He also adds that the exact proof of the result will replace the assumption of soundness (logic does not produce false conclusions from true premises) by another, much weaker, namely ω-consistency;

Gödel's letter to Bernays (April 2); Gödel informally explains some points of his results;

Tarski gives a talk on Gödel's results in Warsaw on April 14 (see Tarski 1930–1931a). It was probably on the first reports about (**TG2**).

Tarski 1931 is published; it contains remarks on the concept of satisfaction and provides the definition of definability, in particular, Tarski remarks that there are undefinable sets of numbers, because the family of definable sets of numbers is denumerable (the set of all sets of numbers is not denumerable). This statement can be understood as the approximation of the statement that the set of all truths about numbers is not definable, but it is only a hypothesis that Tarski was thinking in such a way;

The Zermelo–Gödel correspondence concerning the incompleteness theorem; Gödel observes in the letter of October 12 (see Gödel 2003a, pp. 427–429) that he demonstrated that the set of provable arithmetical statements is a proper subset of true arithmetical statements, because the former is expressible in arithmetic itself, but the latter is not (see below); this correspondence is one the main arguments that Gödel already had (**UT**) at that time;

Carnap reports about the meeting of the Vienna Circle on 1931.07.02 (see Köhler 2002, p. 97) that there are common (*ordenliche*) concepts which are not definable.

1932

Gödel 1932 is published, no (**UT**);

Tarski 1932 is published, a version of (**UT**) is published for languages of the infinite order (in the sense of the theory of logical types; Tarski explicitly says (Tarski 1932, p. 616) that he was influenced by Gödel 1931;

Gödel's letter to Carnap (September 11; see below); Gödel says that he will give a truth-definition in the second part of his work, that is, in a sequel to Gödel 1931 (this work was never written);

Gödel's letter to Carnap, November 28, 1932 (see below, (f) in quotations from Gödel); Gödel mentions Tarski 1932.

1933

Tarski 1933 is published; (**UT**) is formulated for languages based on the theory of logical types in a more detailed way than in Tarski 1932.

1934

The Gödel Princeton lectures (Gödel 1934); Gödel repeats the argument from his letter to Zermelo;

Carnap 1934, Carnap 1934a are published, a version of (**UT**); the fixed-point lemma (**FPL**).

1935

Tarski 1935 is published and preprints distributed; it contains (**UT**) in full generality, important additions and historical comments about the relation of Tarski's results to that of Gödel (see Tarski 1935, pp. 277–278, partially quoted below); Gödel's influence is acknowledged; German translation of Tarski 1933 is frequently dated as published in 1936. However, the volume of *Studia Philosophica* in which this work appeared is officially dated 1935. I follow this indication, although there is some evidence that this volume actually appeared in 1936;
Tarski's letter to Twardowski on July 24, 1935 (unpublished, the original is in the Archive of *Studia Philosophica* in Poznań); Tarski says that his historical comments were "entirely squared with Carnap and Gödel"; the issue was important for Tarski, due to suggestions (probably inseminated by Leśniewski) that he plagiarized Gödel;
Gödel 1935 is published; Gödel points out the similarity between Carnap and Tarski.

1936

Tarski 1935 is published; the (un)definability of truth is noted with reference to Gödel.

1939

Tarski 1939 is published; undecidable sentences via the concept of truth; the first complete demonstration that $Cn\mathbf{AR} \subset \mathbf{VER}(\mathbf{\Omega})$.

8.8.2 A Selection of Quotations from Gödel

(a) Gödel 1931, p. 151:

> The method of proof [of incompleteness – J. W] just explained can clearly be applied to any formal system that, first, when interpreted as representing a system of notions and propositions, has at its disposal sufficient means of expression to define the notions occurring in the argument above (in particular, the notion "provable formula") and in which, second, every provable formula is true in the interpretation considered. The purpose of carrying out the above proof with full precision in what follows is, among other things, to replace the second of the assumptions just mentioned by a purely formal and much weaker one.

(b) Gödel 1931, p. 181, note 48a:

> […] the true reason for the incompleteness inherent in all formal systems of mathematics is that the formation of ever higher types can be continued into the transfinite […] while in any formal system at most denumerably many of them are available. For it can be shown that the undecidable propositions constructed here become decidable whenever appropriate higher types are added. […]. An analogous situation prevails for the axioms system of the set theory.

(c) Gödel' letter to Bernays, April, 1931 (Gödel 2003, p. 97; tr. by S. Feferman; I use slightly different symbolism):

> […] the principle according to which the class $W(x)$ [x is true – J. W.] is defined is recursive; I first define, what W means for the simplest propositions of all (numerical equations, etc.) and then go on to more complicated propositions, say according to the following schema:
>
> $$W(\neg A) = \mathrm{Df.}\ \neg W(A);$$
> $$W(A \vee B) = \mathrm{Df.}\ W(A) \vee W(B);$$
> $$W(\forall v A(v)) = \mathrm{Df.}\ \forall v(W(A(a)), \text{for any substitution of a term } a \text{ for the variable } v).$$
>
> […] the procedure is only vaguely sketched by that. Simultaneously and independently of me (as I gathered from a conversation), Mr. Tarski developed the idea of defining the concept of "true proposition" in this way. […].
>
> As concerns the decidability of the undecidable propositions in higher systems, that results immediately from the properties of the concept $W(x)$.

(d) Gödel's letter to Zermelo, October 12, 1931 (Gödel 2003, pp. 427–429; tr. by J. Dawson):

> This concept [that is, of the correct formula – JW], however, may not, without further ado, be traced back to a combinatorial property of formulas (but rather rests upon the meaning of the symbols), and therefore may not be traced back in arithmetized metamathematics to simple arithmetical concepts; or, in other words: The class of correct formulas in *not* expressible by means of a class sign of the given system [a footnote: More precisely stated, it is of course a question of the class of those *numbers* that are assigned to correct formulas […]. The situation is quite otherwise for the concept "provable formula" […]. In connection with what has been said, one can moreover carry out my proof as follows: The class W of correct formulas *is never* coextensive with a class sign of that same system (for the assumption that that is the case leads to a contradiction). The class B of provable formulas *is* coextensive with a class of that same system […]; consequently B and W cannot be coextensive with each other. But because $B \subset W$, holds, i.e., there is a correct formula A that is not provable. Because A is correct, not-A is also not provable, i.e., A is undecidable. This proof has, however, the disadvantage that it furnishes no construction of the undecidable statement and is not intuitionistically unobjectionable.
>
> I would still like to remark that I see the essential point of my result not in that one can somehow go outside any formal system (that follows already according to the diagonal procedure), but that for every formal system of metamathematics there are statements which are *expressible* within the system but which may *not be decided* from the axioms of that system […]. O course, [my explanations] should not be taken as "proof". The proof is rather to be found at the places of my paper I've cited [that is Gödel 1931 – J. W.)].

(e) Gödel's letter to Carnap, September 11, 1932 (Gödel 2003, p. 347, tr. by W. Goldfarb)

> In my judgment, this error [in Carnap's account of analyticity – J. W.] may only be avoided by regarding the domain of the function variables [that is, the predicate letters – J. W.] not

as the predicates of a definite language, but rather as all sets and relations whatever. On the basis of this idea, in the second part of my work I will give an definition for "truth", and I am of the opinion that the matter may not be done otherwise, and that one can *not* view the higher functional calculus [that is, higher order logic – JW] semantically. That is, one can of course build up a higher functional calculus on a semantic basis […].

(f) Gödel's letter to Carnap, November 28, 1932 (Gödel 2003, pp. 355–357; tr. by W. Goldfarb)

In order to be able to carry out the matter [of defining 'analytic' – J. W.] in general, that is, for functions of arbitrary finite type, one needs a variable of the next higher type (type ω), i.e., one which runs through all finite types. This could be foreseen a priori, since one can never define "analytic" in the same system – otherwise contradictions will result. I believe moreover that the interest of this definition does not lie in the clarification of the concept of "analytic", since one employs in it the concepts "arbitrary sets", etc., which are just as problematic. Rather I formulate it only for the following reason: with the help one can show that undecidable sentences become decidable in systems which ascend farther in the sequences of types. By the way, as you perhaps know, Tarski has given a similar definition for "analytic" in a paper to appear in the Polish language, about which he reported in the Anzeiger der Wiener Akademie 1932, no. 2.

(g) Gödel (1934, p. 363):

So we see that the class α of numbers of true formulas cannot be expressed by a propositional function of our system, whereas the class β of provable formulas can. Hence […] if we assume […] every provable formula is true […] there is a proposition A which is true but not provable. $\sim A$ then is not true and therefore not provable either, i.e., A is undecidable.

(h) Gödel (1934, p. 363) (a footnote added in 1964):

For a closer examination of this fact [of antinomies – J. W.] see A. Tarski' papers [Tarski 1933 and 1944] […]. In these two papers the concept of truth relating to sentences of a language is discussed systematically.

(i) Gödel's lecture on the continuum hypothesis delivered at Brown University in 1940 (Gödel 1995, p. 181)

[…] this metamathematical notion of truth, i.e., that is, the class of numbers of true propositions, can be defined by a method similar to the one which Tarski applied for the system of *Principia Mathematica*.

(j) Gödel's, letter to Arthur Burks, probably in the early 1960s (Burks 1966, p. 55):

> […] a complete epistemological description of a language *A* cannot be given in the same language *A*, because the concept of truth of sentences of *A* cannot be given in the same language *A*, because the concept of truth of sentences of *A* cannot be defined in *A*. It is this theorem which is the true reason for the existence of undecidable propositions in the formal systems containing arithmetic. I did not, however, formulate it explicitly in my paper of 1931 but in my Princeton lectures of 1934. The same theorem was proved by Tarski in his paper on the concept of truth published in 1933 […].

(k) Gödel's letters to Hao Wang, December 7, 1967 (Gödel 2003a, pp. 397–398. pp. 403–404):

> Non-finitary reasoning in mathematics was widely considered to be meaningful only to the extent to which it can be "interpreted" or "justified" in terms of a finitary metamathematics. This view, almost unavoidably, leads to an exclusion of non-finitary reasoning from metamathematics. […] my objectivistic conception of mathematics and metamathematics in general, and of transfinite reasoning in particular, was fundamental also to my other work in logic. […]. […] it should be noted that the heuristic principle of my construction of undecidable number–-theoretic propositions in the formal systems of mathematics is the highly transfinite concept of "objective mathematical truth" as *opposed* to that of "demonstrability" [with] which it was generally confused before my own and Tarski's work.

(k) Gödel's letter to Balas, ca. about 1970, a draft, not sent (Gödel 2003, p. 9–10):

> I have explained the heuristic principle for the construction of propositions undecidable in a given formal system in the lectures I gave in Princeton in 1934. […]. The occasion for comparing truth and demonstrability was an attempt to give a relative model-theoretic consistency proof of analysis in arithmetic. […] Hence, if truth were equivalent to provability, we would have reach our goal. However, (and this is a decisive point) it follows from the correct solution of the semantic paradoxes, that the concept of "truth" of the propositions of a language *cannot be expressed* in the same language, while provability (being an arithmetical relation) *can.* Hence, true $\neq$ provable. […] However, in consequence of the philosophical prejudices of our times: 1. Nobody was looking for a relative consistency proof because it was considered axiomatic that a consistency proof must be finitary in order to make sense; 2. a concept of objective mathematical truth as opposed to demonstrability was viewed with greatest suspicion and widely rejected as meaningless.

8.8.3 Tarski's Historical Explanations

Tarski 1933, p. 152 (present in the Polish original), pp. 277–278 (added in German version):

> The results presented in this paper date for the most part from 1929. I discussed them, in particular, in two lectures given under the title 'On the concept of truth in reference to formalized deductive sciences' at the Logic Section of the Philosophical Society in Warsaw (October 8, 1930) and at the Polish Philosophical Society in Lwów (December 5, 1930). A short report of these lectures is in Tarski [1930–31]. […] For reasons beyond my control, publication was delayed by two years. In the meantime the original text was supplemented by some substantial additions (see p. 247) […] a summary of the chief results of the paper was published in Tarski [1932].
>
> […]
>
> I may say quite generally that all my methods and results, with the exception of those at places where I have expressly emphasized this – [here are references to Leśniewski's treatment of semantic antinomies and Gödel's method of arithmetization – J. W.) – were obtained by me quite independently. […]
>
> I should like to emphasize the independence of my investigations regarding the following points of detail: (1) the general formulation of the problem of defining truth […]; (2) the positive solution of the problem, i.e. the definition of the concept of truth for the case where the means available in the metalanguage are sufficiently rich […]; (3) the method of proving consistency on the basis of the definition of truth […]; (4) the axiomatic construction of the metasystem […]; (5) the discussions […] on the interpretation of the metasystem in arithmetic, which already contain the so-called 'method of arithmetizing the metalanguage' which was developed far more completely and quite independently by Gödel. Moreover, I should like to draw attention to results not relating to the concept of truth but to another semantical concept, that of definability […].
>
> In the one place in which my work is connected with the ideas of Gödel – in the negative solution of the problem of the definition of truth for the case where the metalanguage is not richer than the language investigated – I have naturally expressly emphasized this fact […]; it may be mentioned that the result so reached, which very much completed my work, was the only one subsequently added to the otherwise already finished investigations.

Tarski 1939, p. 561:

> It is my intention in this paper to add somewhat to the observations already made in my earlier publications on the existence of undecidable statements in systems of logic possessing rules of inference of a "non-finitary" ("non-constructive") character.[…]. I also wish to emphasize once more the part played by the concept of truth in relation to problems of this nature. […]. At the end of this paper I shall give a result which was not touched upon in my earlier publications. It seems to be of interest for the reason […] that it is an example of a result obtained by a fruitful combination of the method of constructing undecidable statements (due to K. Gödel) with the results obtained in the theory of truth.

8.8.4 Comments

Who discovered (**UT**)? Although Carnap's contributions were essential, Gödel and Tarski appear as the main competitors. The name 'the Tarski undefinability theorem' suggests that the discovery should be credited to Tarski (the label 'the Gödel–Tarski theorem' sometimes appears, for instance, in Smoryński 1985, p. 6). Tarski's way to this theorem was long—from Tarski 1930 through Tarski 1933, its German translation in 1935, Tarski 1939 to Tarski, Mostowski, Robinson 1953. We do not know when Tarski established (**UT**) as a separate result. It is certain that he did not have it in 1930, because no mention of it occurs in Tarski 1930–1931. The first announcement appeared in Tarski 1932. Tarski gradually extended his results. At the beginning, he wanted to achieve a truth-definition free of semantic antinomies. Clearly, such a definition could not be formulated for natural language, because it confused the object language and the metalanguage. On the other hand, the thesis that truth is not definable for natural language, because it would produce antinomies, cannot be considered as a version of (**UT**), although it perhaps suggested a heuristics to proceeding further. The first version of (**UT**) appeared in Tarski 1933 as the theorem that truth-definition cannot be formulated for languages of infinite order (in other words, in one formula for all logical types). He improved these results (see Tarski's quoted historical explanations) under Gödel's influence. The method of arithmetization allowed him to see in which sense **ML** must be richer than **L** in order to suffice for **SDT**. Consequently, truth for **L** cannot be defined in **L**, but requires **ML**—stronger than the object-language. Moreover, Tarski abandoned his earlier hope that **SDT** can provide a proof of the consistency of the theory of types and, *a fortiori*, arithmetic. Further Tarski's subsequent works explored the semantic method of proving incompleteness. For Tarski, this method was satisfactory, although non-constructive. We can summarize Tarski's way to (**UT**) as follows. This theorem combines: (a) the semantic definition of truth; (b) the statement that this definition assumes some conditions for its formalization, particularly, that **ML** is richer than **L**; (c) if these conditions are not fulfilled the concept of truth for **L** is not definable in **L**. In this context, the method of arithmetization appears as an auxiliary device. The diagonalization lemma is much more important, because it shows that the negation of at least one **T**-equivalence is produced by a truth-definition for **L** in this language. We have the following sequence of steps. Formalize a truth-definition in a language **L**. Assume that there is a formula of **L** which provides a definition of truth in the sense of the concept of definability. If this definition is materially adequate, it should yield all **T**-equivalences. Use (**FPL**). If it leads to a contradiction, that means that **L** cannot formalize the proposed truth-definition. To return to (**TT^3**) (it is implicitly in Tarski, Mostowski and Robinson 1953), we have a dilemma of either defining the diagonal function or defining truth. Because the diagonal function is a natural construction, the **T**-predicate is not definable.

As I remarked earlier, the discovery of the Gödel–Zermelo correspondence (see Grattan–Guiness 1979) changed the received (circa, 1975) historical perspective to the claim that (**UT**) was established by Gödel, as early as in 1931, that is, before Tarski (other relevant works in this context are Murawski 1998, Feferman 1999, Murawski 1999a). I will argue that these opinions should be taken with a measure of reserve. I will mainly concentrate on (**UT**) (including the heuristics leading to this theorem), but I neglect comments concerning the historical accuracy of the passages quoted in subsections **B** and **C**. It is a pity that we do not know the content of discussions between Carnap, Gödel and Tarski in Vienna in 1930, and perhaps also in 1932. The only historical hint is to be find in (c). Gödel surely refers here to conversations in Vienna in 1930. Although we know that Carnap, Gödel and Tarski met together at least once, but what they talked about remains unknown. In this situation, all speculations are pointless.

I outlined Tarski's path above. What about Gödel's? He insisted that his objectivistic conception of mathematical truth was a guide for the undecidability theorems. Yet this view is not discernible in Gödel 1929, Gödel 1930. One can even get the impression that Gödel was dissatisfied with the non-constructive character of his completeness proof. In Gödel 1931 we find a remark about the similarity of the sentence (*) 'I am not provable' (the crucial element in Gödel's argument) to the Liar sentence and the Richard paradox. Gödel adds (p. 149, footnote 14) that "any epistemological [that is, semantic—J. W.] antinomy could be used for a similar proof of the existence of undecidable propositions". It is a strange assertion, because the Liar sentence leads to inconsistency, but (*) to undecidability. Clearly, Tarski's heuristics used semantic antinomies as a sign of unprovability, not undecidability. Gödel insists very strongly that semantic notions should be replaced by syntactic [constructive, intuitionistically acceptable concepts (see quotation (a)]. Nothing like that occurs in Tarski. Quotation (b) adds something new, namely that the formation of new (higher) types can be continued into the transfinite, although in any formal system only a finite number of types is available. No remark on truth as an infinitary concept occurs in Gödel 1931.

Quotations (b) and (c) are crucial. The former contains a definition of truth and a remark that "the decidability of the undecidable propositions in higher systems, that results immediately from the properties of the concept of $W(x)$". Looking at the proposed definition in (b), it contains no hint on how to define truth for atomic propositions. Moreover (see Feferman 2003, p. 45, note m), this definition can only be applied to languages in which every object has a name, that is, associated with canonical models. Thus, Gödel's truth-definition is actually less general than Tarski's construction via the concept of satisfaction. Consequently, one should beware of Gödel's proclamation that Tarski developed the idea of defining the concept of true proposition in "this way". For Gödel, the definition of truth was an auxiliary device which allows to show nothing more than that the undecidabilities within a given system, become decidable (provable) in its metasystem. This explains the shape of Gödel's definition in the letter to Bernays. In other words, he did exactly what he needed to do in order to show how one can decide (in the metasystem) the undecidable sentences from the system. Yet we do not know how to derive, in higher system, the decidability of the undecidable propositions in a higher system, unless we stay with the trivial observation that they are added on as new axioms. At any rate, it seems that Gödel was much more interested in provability in higher systems than the (un)definability of truth. The letter to Zermelo became the main reason to maintain that Gödel had (**UT**). In fact, he says that because $B \subset W$ holds there are unprovable correct formulas. However, it is clear from his further explanations that (un)decidability is the main target here, and not the question whether truth is definable or not. Gödel repeats his standing view (at that time) that "this proof has, however, the disadvantage that it furnishes no construction of the undecidable statement and is not intuitionistically unobjectionable". At the end of the letter, Gödel says that his explanations do not constitute a proof, which can be "rather" found in Gödel 1931. However, the demonstration that $B \subset W$ is not to be found anywhere in this famous paper.

Clearly, the formula $B \subset W$ would be considered as (**UT**) in light of the further development of metamathematics. Gödel's later explanations well confirm this suggestion. However, what Gödel established in 1931 (it is explicit in the letters to Bernays and Zermelo) and repeated in 1934 is that the set of provable sentences (of arithmetic and its oversystems) is a proper subset of the set of true sentences. From this he concluded that there exist arithmetical sentences that are simultaneously true and unprovable. This was shown explicitly in the letter to Zermelo and in the Princeton lectures. One can say that his informal argument for the incompleteness of arithmetic (present in Gödel 1931) implicitly suggests the same. Is this result—that is, the first incompleteness theorem in its semantic version—equivalent to (**UT**)? The way out of considering Gödel's statements as a formulation of (**UT**) consists in identifying 'being expressible in **Th**' with 'being definable in **Th**'; Feferman takes this path (see Feferman 1984, pp. 158–159). An indirect suggestion for that can be derived from Gödel's letter to Bernays where he says (p. 95) that the class of numbers of true (correct) formulas can be defined in the metasystem. Combining both letters, we can conclude that such a definition is impossible within the system. Gödel even established (**FPL**), and used it for an informal demonstration that undecidable sentences of the system are decidable in the metasystem, although he was thinking more about their provability than their truth. However, Gödel never used (**FPL**) to establish the undefinability of truth; at least he did not do so explicitly. Nobody can deny that Gödel was on the right track to (**UT**), but he did not complete the job. Hence, it is incorrect to say, without further comments, histoprical as well as substantial that (**UT**) is "sometimes presente as an easy companion to Gödel's first incompleteness theorem" (Ray 2018, p. 709). In fact, this theorem is sometimes presented in such a way (not by Greg Ray to be fair), but such a view seems to be unfair. A more accurate opinion is expressed by Raymond Smullyan (see Smullyan 1992, Chapter II; Smullyan 2001, pp. 78–79) who considers **UT** as of a special interest. In particular, according to Smullyan, Tarski's method immediately leads to (semantic) proof of (**TG2**), but original Gödel's proof does not suffice for proving the undefinability result.

For Tarski, the fundamental problem consisted in constructing a definition that would clarify the concept of truth. Tarski had a general truth-definition, but his (**UT**) shows that the concept of truth is not definable in some cases, namely in **AR** and stronger theories. Gödel had no such truth-definition and only considered the problem of the expressibility (or definability) of the set of true numbers in **AR** (the Gödel numbers of true sentences of **AR**), and found that this set is not arithmetical. In other words, Gödel was not interested in the positive problem of truth-definition (the Murawski 1998 for the importance of the difference between, Gödel's so to speak, the local approach to truth versus Tarski's global point of view. Tarski's approach afforded him the possibility of formulating (**UT**) as a metamathematical theorem and of proving it by (**FPL**). I am not sure whether in 1931 or even 1934, Gödel himself thought about the undefinability or non-expressibility of truth in this way. Thus, I cannot agree with Hao Wang's opinion (see Wang 1986, p. 144) that "on the essential points Gödel had not only anticipated Tarski but also understood better what was involved". My assessment of the historical situation relater to **(TG2)** and (**TT**) is closely related to Smullyan's perspective (see comments in Sect. 8.5). Yet it is not my intention to suggest which result is more important.

There still remains an intriguing problem of how Gödel was so insensitive to the issue of a general formal truth-definition (see also Von Plato 2019 for collecting together various, published and unpublished), Gödel's remarks on truth expressed in 1929–1931). It is not clear why he resigned from completing the work mentioned in his 1932 letter to Carnap as far as a truth-definition is concerned. Although the problem concerning Gödel and truth, as well as his sensitivity to philosophical problems, is extensively investigated in many works (see i. a. Feferman 1984, Wang 1986, Wang 1987, Wang 1996, Dawson 1997, Buldt 2002, Krajewski 2003, Von Plato 2019), something needs to be added. Although Gödel's interests in Platonism, Kant, Leibniz and Husserl are well known, it seems that a greater stress should be laid on Gödel's understanding of truth as transcendental (perhaps in Kant's sense). As a result, although he accepted non-finitary reasoning, he did this privately, so to speak, but officially worked inside finitary metamathematics in which there was no room for a general formal theory of truth. For this reason, the concept of truth as transcendental was not captured formally and, as Hintikka said (Hintikka 1997a), was ineffable via formal tools. Gödels's statement, to quote it once again, "This proof has, however, the disadvantage that it furnishes no construction of the undecidable statement and is not intuitionistically unobjectionable", expresses precisely his foundational scruples. Most commentators say that his caution (to use Feferman's phrase once again; see Feferman 1984) concerning the concept of truth resulted partly as an effect of the influence of his environment. He worked on the Hilbert program, and his proof of the completeness theorem was an important step on the way to finitary metamathematics. The incompleteness theorems went in the opposite direction and this produced a tension between acceptable (constructive) forms of provability and the not quite desirable transfinite nature of truth. Even if Gödel changed his views to more officially sympathetic toward infinitary tools, he regarded the concept of truth as not tractable via exact mathematical tools. Yet he was ready to explain why that was so. To sum up briefly, Tarski considered the concept of truth as mathematically definable, although under some additional constraints (some of them were suggested by Gödel), whereas Gödel's view was rather that this notion transcends mathematical tools—at least if we investigate **AR** and its extensions. Thus, (**FC**) as the fundamental conclusion concerning the syntax-semantics relation could be stated by Gödel only informally —but formally by Tarski. Ironically and perhaps even paradoxically, Gödel's philosophical views were closer to (**UT**) than those of Tarski. It is a good lesson about the role of philosophical context in doing mathematics.

References

Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E., Romanowska, A., Woleński, J. (Eds.) (2004). *Provinces of logic determined. Essays in the memory of Alfred Tarski. Parts VI, IV and VI* (*Annals of Pure and Applied Logic*, 127). Amsterdam: Elsevier.

Barwise, J. (1975). *Admissible Sets and Sructures*. Berlin: Springer.

Barwise, K., Perry, J. (1983). *Situations and Attitudes*. Cambridge, Mass: The MIT Press.

Beeh, V. (2003). *Die halbe Wahrheit. Tarskis Dfinition & Tarskis Theorem*. Paderborn: Mentis.

Bell, J., Machover, M. (1977). *A Course in Mathematical Logic*. Amsterdam: North-Holland.

Buldt, B. (2002). Philosophische Implikationen der Gödelsche Sätze? Ein Bericht. In Köhler et al. (Vol. 2, 395–438).

Burks, A. (1966). Editior's comments. In Von Neumann (1966), 53–56.

Carnap, R. (1934). *Logische Syntax der Sprache.* Wien: Springer; Engl. tr., London: Routledge and Kegan Paul 1937.

Carnap, R. (1934a). Die Antinomien und die Unvollständingkeit der Mathematik. *Monatshefte für Mathematik und Physik, 41,* 263–284.

Carnap, R. (1947). *Meaning and Necessity. A Study in Semantics of Modal Logic*. Chicago: University of Chicago Press; 2nd ed. 1956.

Casari, E. (2006). *La matematica della verità. Strumenti matematici della semantica logica*. Torino: Bollati.

Chang, C. C., Keisler, J. (1990). *Model Theory*. Amsterdam: North-Holland.

Cook, R. T. (2014). *The Yablo paradox. An Essays on Circularity*. Oxford: Oxford University Press.

Davis, M. (Ed.). (1965). *The Undecidable. Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions*. Hewlett, NY: Raven Press.

Dreben, B., Van Heijenoort, J. (1986). Introductory note. In Gödel (1986), 44–59.

Enayat, A.., Kossak, R. (Eds.) (2003). *Nonstandard Models of Arithmetic and Set Theory*. Providence, R.I.: American Mathematical Society.

Enderton, H. B. (1972). *A Mathematical Introduction to Logic*. New York: Academic Press.

Feferman, S. (1984). Kurt Gödel: Conviction and caution. *Philosophia Naturalis*, 21, 546–562; repr. in Feferman (1998), 150–164.

Feferman, S. (1989). *The Number Systems. Foundations of Algebra and Analysis*. New York: Chelsea Publishing Company.

Feferman, S. (1998). *In the Light of Logic*. Oxford: Oxford University Press.

Feferman, S. (1999). Tarski and Gödel: Between the lines. In Woleński, Köhler (1999), 53–63.

Feferman, S. (2003). Introductory note. In Gödel (2003), 43–78.

Gaifman, Ch. (2003). Non-standard models in a broader perspective. In Enayat, Kossak (2003), 1–22.

Glanzberg, M. (Ed.) (2018). *The Oxford Handbook of Truth*. Oxford University Press: Oxford.

Goble, L. (Ed.) (2001). *The Blackwell Guide to Philosophical Logic*. Oxford: Blackwell.

Gödel, K. (1929). *Über die Vollständigkeit des Logikkalkül*, Wien Universität, repr. with Eng. tr. in Gödel (1986), 60–101.

Gödel, K. (1930). Die Vollständigkeit der Axiome des logischen Funktionenkalkül. *Monatshefte für Mathematik und Physik*, 37 349–360; repr. with Eng. tr. in Gödel (1986), 102–123.

Gödel, K. (1930a). Einige metamathematische Resultate über Entscheidungsdefinitheit und Widerspruchsfreiheit. *Anzeiger der Akademie der Wissenschaften in Wien*, 67, 214–215; repr. with Eng. tr. in Gödel (1986), 140–143.

Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38, 173–198; repr. with Eng. tr. in Gödel (1986), 145–195.

Gödel, K. (1931a). Natrag. *Erkenntnis*, 2, 149–150; repr. with Eng. tr. in Gödel (1986), 202–204.

Gödel, K. (1932). Über Vollsändigkeit und Widerspruchsfreiheit. *Ergebnisse einer mathematischen Kolloqiums*. 3, 20–21; repr. with Eng. tr. in Gödel (1986), 234–237.

Gödel, K. (1934). *On Undecidable Propositions of Formal Mathematical Systems* (mimeographed notes by S. C. Kleene and J. B. Rosser). Princeton: Princeton; repr. (with additions) in Gödel (1986), 346–371.

Gödel, K. (1935). Review of Carnap (1934). *Zentralblatt für Mathematik und ihre Grenzgebiete*, 11, 1; repr. with Eng. tr. in Gödel (1986), 388–389.

Gödel, K. (1986). *Collected Works*, v. I. Oxford: Oxford University Press.

Gödel, K. (1995). *Collected wWorks*, v. III: *Unpublished Essays and Lectures*. Oxford: Oxford University Press.

Gödel, K. (2003). *Collected works*, v. IV: *Correspondence A–G*. Oxford: Clarendon Press.

Gödel, K. (2003a). *Collected works*, v. V: *Correspondence H–Z*. Oxford: Clarendon Press.

Goldstern, M., Judah, H. (1995). *The Incompleteness Phenomenon. A New Course inMathematical Logic*. Wellesley: A K Peters.

Gómez-Torrente, M. (2004). The indefinability of truth in the "Wahrheitsbegriff". In Z. Adamowicz, S. Artemov, D. Niwiński, E. Orłowska, A. Romanowska, J. Woleński (2004), 27–38.

Grattan-Guiness, I. (1979). In memoriam Kurt Gödel: His 1931 correspondence with Zermelo on his incompletability theorem. *Historia Mathematica*, 6, 294–304.

Gries, D., Schneider, F. B. (1993). *A Logical Approach to Discrete Mathematics*. Berlin: Springer.

Grzegorczyk, A. (1974): *An Outline of Mathematical Logic*. Dodrecht: D. Reidel.

Hàjek, P., Pudlák, P. (1998). *Metamathematics of First-Order Arithmetic*. Berlin: Springer.

Henkin, L. (1949). A proof of completeness for the first-order functional calculus. *The Journal of Symbolic Logic*, 14, 159–166.

Hintikka, J. (1991). *Defining Truth, the Whole Truth and Nothing but the Truth*. Reports from the Department of Philosophy. Helsinki: University of Helsinki, no. 2; repr. in Hintikka (1997), 46–103.

Hintikka, J. (1996). *The Principles of Mathematics Revisited.* Cambridge: Cambridge University Press.

Hintikka, J. (1997). *Lingua Universalis vs. Calculus Ratiocinator: An Ultimate Presupposition of Twentieth-Century Philosophy*. Dordrecht: Kluwer.

Hintikka, J. (1998). *Language, Truth and Logic in Mathematics*. Dordrecht: Kluwer.

Kaye, R. (1991). *Models of Peano Arithmetic*. Oxford: Clarendon Press.

Köhler, E. (2002). Gödel und Wiener Kreis. In Köhler et al. (2002) v. 1, 83–108.

Köhler, E. et al. (Eds.) (2002). *Kurt Gödel. Wahrheit & Beweisbarkeit*, Band 1.: *Dokumente und historische Analysen*; Band 2.: *Kompendium zu Werk*. Wien: öbv et htp

Kokoszyńska, M. (1936). Syntax, Semantik und Wissenschaftlogik. In *Actes du Congès International de Philosophie Scientifique*, v. 3. Paris: Herman, 9–14; repr. in Pearce, Woleński (1988), 271–275.

Kossak, R. (2004). Undefinability of truth and non-standard models. In Adamowicz, Artemov, Niwiński, Orłowska, Romanowska, Woleński (2004), 115–123.

Kotlarski, H. (2004). The incompleteness theorems after 70 years. In Adamowicz, Artemov, Niwiński, Orłowska, Romanowska, Woleński (2004), 125–138.

Krajewski, S. (2003). *Twierdzenie Gödla i jego interpretacje filozoficzne. Od mechanicyzmu do postmodernizmu* (Gödel's Theorem and Its Philosophical Interpretations. From Mechanicism to Post-Modernism). Warszawa: Wydawnictwo Instytutu Filozofii i Socjologii PAN.

Lindström, P. (2002). *Aspects of Incompleteness*. Natick, Mass.: A K Peters.

Mendelson, E. (1997). *Introduction to Mathematical Logic*. London: Champan & Hall.

Menger, K. (1994). *Reminiscences of the Vienna Circle and the Mathematical Colloqium*. Dordrecht: Kluwer.

Murawski, R. (1998). Undefinability of truth. The problem of priority: Tarski vs. Gödel. *History and Philosophy of Logic,* 19, 153–160.

Murawski, R. (1999). *Recursive Functions and Metamathematics. Problems of Completeness and Decidability*. Dordrecht: Kluwer.

Murawski, R. (1999a). Undefinability vs. definability of satisfaction and truth. In Woleński, Köhler (1999), 203–215.

Pantsar, M. (2009). *Truth, Proof and Gödelian Arguments. A Defence of Tarskian Truth inMathematics*. Helsinki: University of Helsinki.

Pearce, D., Woleński, J. (Eds.) (1988). *Logische Rationalismus. Philosophische Schriften der Lemberg–Warschauer Schule*. Frankfurt a. M.: Athenäum.

Poizat, B. (2000). *A Course in Model Theory. An Introduction to Contemporary Mathematical Logic*. Berlin: Springer.

Popper, K. (1955). A note on Tarski's definition of truth. *Mind*, 64, 388–391; repr. in Popper (1972), 335–340.

Popper, K. (1972). Objective Knowledge. An Evolutionary Approach. Oxford: Clarendon Press; repr. with additions, 1974.

Ray, G. (2018). Tarski on the concept of truth. In Glanzberg (2018), 695–719.

Rojszczak, A., Cachro, J., Kurczewski, G. (Eds.). (2003). *Dimensions of Logic and Science. Selected Contributed Papers from the 11th International Congress of Logic, Methodology, and Philosophy of Science*. Kraków, 1999. Dordrecht: Kluwer.

Rosser, J. B. (1936). Extensions of some theorems of Gödel and Church. *The Journal of Symbolic Logic*, 1, 87–91; repr. in Davis (1965), 231–235.

Sirakov, B., Ney de Souza, P., Viana, M. (Eds.) (2019). *Proceedings of International Congress of Mathematicians. Rio de Janeiro 2018*, v. 3. Singapore: World Scientific.

Smorynski, C. (1985). *Self-Reference and Modal logic*. Berlin: Springer.

Smullyan, R. M. (1992). *Gödel's Incompleteness Theorem*. Oxford: Oxford University Press.

Smullyan, R. M. (1994). *Gödel's Incompleteness Theorem*. Oxford: Oxford University Press.

Smullyan, R. M. (2001). Gödel's incompleteness theorems. In Goble (2001), 72–89.

Tarski, A. (1930–1931). O pojęciu prawdy w odniesieniu do sformalizowanych nauk dedukcyjnych (On the Concept of Truth in Reference to Formalized Deductive Sciences). *Ruch Filozoficzny*, 12, 210–211; repr. in Tarski (1986), v. 4, 559.

Tarski, A. (1930–1931a). O pewnym systemie logiki matematycznej i wynikających z niego zagadnieniach metodologicznych (On certain system of mathematical logic and the resulting methological issues). *Ruch Filozoficzny*, 12, 222; Eng. tr. (by J. Tarski) in Woleński (2004c), 123.

Tarski A. (1931). Sur les ensembles définissables de nombres réels. I. *Fundamenta Mathematicae*, XVII, 210–239; Eng. tr. in A. Tarski (1956), 110–142.

Tarski, A. (1932). Der Wahrheitsbegriff in der Sprachen der deduktiven Wissenschaften. *Akademie der Wissenschaften in Wien. Matematisch-Wissenschaftliche Anzeiger*, 69, 23–25; repr. in A. Tarski (1986), v. 1, 613–617.

Tarski, A. (1933). *Pojęcie prawdy w językach nauk dedukcyjnych* (The concept of truth in languages of deductive sciences). Warszawa: Towarzystwo Naukowe Warszawskie; Germ. tr. (with additions) as Tarski (1935).

Tarski, A. (1933a). Einige Betrachtungen über die Begriffe ω-Widerspruchsfreiheit und der ω-Vollständigkeit. *Monatshefte für Mathematik un Physik*, 40, 97–112; repr. in Tarski (1956), 279–295.

Tarski, A. (1935). *Der Wahrheitsbegriff in den formalisierten Sprachen. Studia Philosophica*, 1, 261–405; repr. in Tarski (1986), v. 2, 51–198. Engl. tr. (*The Concept of Truth in Formalized Languages*) in Tarski (1956), 152–278 (page-references to Tarski (1933)).

Tarski, A. (1939). On undecidable sentences in enlarged systems of logic and the concept of truth. *The Journal of Symbolic Logic*, 4(3), 105–112; repr. in Tarski (1986), v. 2, 559–570.

Tarski, A. (1956). Logic, Semantics, Metamathematics. Papers of 1923 to 1938 (tr. by J. H. Woodger). Oxford: Clarendon Press; 2nd ed., Indianapolis: Hackett Publishing Company 1983.

Tarski, A. (1969). Truth and proof. *Scientific American*, 220(6), 63–77; repr. in Tarski (1986), v. 4, 399–423.

Tarski, A. (2001). Odczyt Alfreda Tarskiego na Konferencji o Problemach Matematyki w Princeton, 17 grudnia 1946 (Alfred Tarski's lecture at the conference on problems of mathematics, Princeton, November 17, 1946). In Tarski (2001), 396–413.

Tarski, A., Mostowski, A., Robinson, R. M. (1953). *Undecidable Theories*. Amsterdam: Norh-Holland.

Tarski, A. (1986). Collected papers, v. 4. Basel: Birkhäuser; 2nd ed. 2018.

Von Neumann, J. (1966). *Theory of Self-Reproducing Automata*. Urbana: University of Illinois Press.

Von Plato, J. (2019). In search of the source of incompleteness. In Sirakov, Ney de Souza, Viana, v. 3, 4043–4060.

Wang, H. (1986). *Beyond Analytic Philosophy. Doing Justice to What we Know*. Cambridge, Mass: The MIT Press.

Wang, H. (1987). *Reflections on Kurt Gödel*. Cambridge, Mass: The MIT Press.

Wang, H. (1996). *A Logical Journey. From Gödel to Philosophy*. Cambridge, Mass: The MIT Press.

Woleński, J. (2003). Truth and satisfaction by the empty sequence. In Rojszczak, Cachro, Kurczewski (2003), 267–276; repr. in Woleński (2011), 17–24.

Woleński, J. (2004). Gödel, Tarski and truth. *Revue Internationale de Philosophie*, 59, 459–490; repr. in Woleński (2011), 101–123.

Woleński, J. (2006). A generalization of Hume's thesis. *Logica Yearbook 2005*, 215–223; repr. in Woleński (2011), 155–162.

Woleński, J. (2006a). Tarskian and post–Tarskian truth. In Auxier, Hahn (2006), 647–672: repr. in Woleński (2011), 163–182.

Woleński, J. (2010). Truth and consistency. Axiomathes, 20(2), 347–355; repr. in Woleński (2011), 303–311.

Woleński, J. (2011). *Essays on Logic and Its Applications in Philosophy*. Frankfurt a. M.: Peter Lang.

Woleński, J., Köhler, E. (Eds.) (1999). *Alfred Tarski and the Vienna circle. Austro-Polish Connection in Logical Empiricism*. Dordrecht: Kluwer.

Wolf, R. S. (2005). *A Tour through Mathematical Logic*. Washington: The Mathematical Association of America.

Chapter 9
Interpretations, Comparisons and Philosophical Issues

Abstract The last chapter mostly discusses philosophical aspects. According to my general view on **STT**, I elaborate its various aspects and defend this theory against some philosophical objections. The list of the discussed problems is as follows: **STT** as a correspondence theory of truth, the status of **T**-equivalences, truth and meaning, the relative of absolute character of truth as semantically defined, truth and science, comparison of **STT** with minimalism and coherentism, truth and realism, and applications of **STT** to the Gettier problem and determinism.

9.1 Introduction

STT as a (meta)mathematical construction can be regarded as well-established, although there are more or less controversial problems, for instance, the question of its application outside classical model theory, or how to achieve partial **T**-predicates embeddable into **L** without causing semantic antinomies (see Button, Walsh 2018 for an extensive discussion of relations between formal semantics, model theory and philosophy). Some mathematical logicians consider **SDT** as trivial (see Hao Wang's opinion quoted in the Introduction). Perhaps the most negative assessment occurs in Girard 1999, p. 220; see also Girard 2011, pp. 213, 491:

> […] Tarskism. […] there is popular prejudice saying that there is something in Tarski's notion of truth. In fact the notion of truth: meaning might be located in the standard […] interpretation. […]. To understand what is wrong here, let's have a look at a neighboring area: there might still be linguists which explain the French sentence «*Guillame est étudiant*» by means of somebody called Bill, whom happens to be a student….what a brilliant idea. […]. In fact the notion of truth à la Tarski […] avoid complete triviality by the use of magical expression «meta»; we presuppose the existence of the meta-world in which logical operations already make a sense; the world discourse can be interpreted in the meta-world, typically the truth of *A* become «meta *A*», and we can in turn explain «meta *A*» by «meta-meta *A*». […]. We are facing a *transcendental* explanation of logic. «*The rules of logic have been given to us by Tarski, who in turn got them from Mr. Metatarski*» something like «*Physical particles act in this way because they must obey the laws of physics*». […]. The abuse of the expression «meta» has completely distorted the relation of logicians to their own field… What to think (of educated) logicians who speak of «truth in the standard model» instead of plain truth (analyse the expression «standard» or «intuitive integers» to mean plain integers… not to speak of this habit of concluding a completely trivial construction by «But you know, it is meta», the universally accepted excuse for the want of an idea. We do not want to question the technical value of the distinction expressed

J. Woleński, *Semantics and Truth*, Logic, Epistemology, and the Unity of Science 45, https://doi.org/10.1007/978-3-030-24536-8_9

> by «meta», which is useful, but only its depth […]: if everything useful were important, then we would spend our life speaking of *soap* and other hygenic artefacts.

This text could also be quoted in other places of this book, for instance, on the occasion of the **L/ML** distinction (see Chap. 7, Sect. 7.4), the discussion of standard models (Chap. 8, Sect. 8.6), or in a general evaluation of **STT** as a philosophical enterprise. I consider Mr. Girard's remarks as a confession of a professional mathematical logician who believes that in using the term 'the set of sentences' no appeal to set theory occurs. I noted in Chap. 7, Sect. 7.4 that explaining the **L/ML** distinction via employing natural languages is only an approximation. Mr. Girard seems to accept a very simplified theory of meaning, according to which saying that *A* in the specialized—for instance, mathematical discourse—means *A*. If the question arises as to how Mr. Girard knows about the meaning of *A*, the probable answer would be "Well, I know that from Mr. Girard". What a brilliant idea. Stopping further polemics with Mr. Girard's parody of Mr. Tarski and Mr. Metatarski, I am inclined to share Wilfrid Hodges' view (see Introduction) that our grandsons and granddaughters will likely to study when they settle down to Tarski's definition of truth for formalized languages in their classes "at 9:30 after school assembly".

The issue of **STT** as a philosophical theory is much more complicated (see Woleński 1999b, Woleński 2018b for a preliminary account, extended in this chapter). I know of only one philosopher that made every similar evaluations of **STT** as Jean-Yves Girard did. It was Neurath who wrote to in letters to Carnap (see Cat & Tuboly, forthcoming):

> (January 15, 1943)
>
> The Scholasticism created Brentanism, Brentano begot Twardowski, Twardowski beget Kotarbiński, Łukasiewicz […], both together begot TARSKI etc., and they are God fathers of OUR Carnap too.
>
> (April 1, 1944)
>
> I like to be in harmony with you. I have the feeling to continue your Logical Syntax period […] before you became Tarskized.

Neurath very strongly objected to semantics, accusing it of introducing absolutism (in understanding of truth) and metaphysical realism into philosophy, views that are at odds with scientific empiricism. He, like as Girard, qualified Tarski's theory of truth as trivial (see Girard 2011, p. 491/492; he compares Tarski's definition with Molière''s saying that opium is the cause of sleeping, because there is a dormitive power in it. However, almost all analytic philosophers and/or philosophical logicians "have the feeling" that **STT** is philosophically interesting—perhaps erroneous, but deserving of serious discussion.

My view is that even if we regard **STT** as a formal construction, (meta)mathematically legitimate, there is a host of philosophical problems associated with Tarski's theory. This chapter discusses several philosophical issues around **STT**; I shall also address to various objections against **STT** (see also Tarski 1944, Stegmüller 1957, Udemadu 1995). Some of these issues were touched on in the preceding chapters. I mean the problem of truth-bearers (Chap. 4, Sect. 4.2), the problem of the epistemological nature of **SDT** (Chap. 7, Sect. 7.5), compositionality (Chap. 5, Sect. 5.2.1), or the relation of truth and logic. Thus, I take for granted that **STT** ascribes 'true' to meaningful sentences, is epistemological, compositional and implies (**BI**) (the principle of bivalence). Other settled matters (see Chap. 7, Sect. 7.2) posit that **STT** has a partial application to natural languages, **SDT** defines the **T**-predicate for **L** (first-order) in **ML** (enough for weak second-order logic), that even through this definition is extensional (it defines the class of truths in a model), the adjective 'true' has a definite intension, and that **STT** satisfies the Russell conditions (sentences, that is items syntactically similar to judgments are truth-bearers), truth is a relation between truth-bearers and something else, truth-definition provides the definition of 'false'). I shall only marginally return to these conclusions, if at all. The content of this chapter is as follows. Section 9.2 intends to show in which sense **SDT** can be qualify as a classical and/or correspondence theory; I also make some remarks about the truth-criterion in the light of **SDT** and the problem of how this definition replies to Brentano's objections (see Chap. 3, Sect. 3.5). The following section is devoted to the status of **T**-equivalences; Sect. 9.3 discusses the status of **T**-equivalences; Sect. 9.4—the relation between truth and meaning. Then (Sect. 9.5), I pass to the question of whether **SDT** is absolute or relative. Section 9.6 discusses the applicability of **STT** to empirical science. Sections 9.7–9.9 compare **SDT** with deflationism and the coherence theory. The last section focuses on **STT** and realism.

9.2 Is STT a Correspondence Theory?

In the Introduction, I quoted Tarski's statement that "the construction of the definition of true sentence and establishing the scientific foundations of the theory of truth—belongs to the theory of knowledge and forms one of the chief problems of philosophy." This opinion is reconfirmed by the following quotations ((a) Tarski 1933, p. 153, p. 155; (b) Tarski 1944, pp. 666–667; (c) Tarski 1969, p. 402/403).

> (a) […] throughout this work I shall be concerned exclusively with grasping the intentions which are contained in the so-called *classical* conception of truth ('true – corresponding with reality') in contrast, for example, with the *utilitarian* conception ('true – in certain respects useful'). […]
>
> Among the manifold efforts which the construction of a correct definition of truth for the sentences of colloquial language has called forth, perhaps the most natural is the search for a *semantical definition.* By this I mean a definition which we can express in the following words:
>
> […] *a true sentence is one which says that the state of affairs is so and so, and the state of affairs is indeed so and so.*
>
> From the point of view of formal correctness, clarity and freedom from ambiguity of the expressions occurring in it, the above formulation obviously leaves much to be desired.

Nevertheless its intuitive meaning is and general intention seems to be quite clear and intelligible. To make this intuition more definite, and to give it a correct form, is precisely the task of a semantical definition.

(b) We should like our definition to do justice to the intuitions which adhere to the *classical Aristotelian conception of truth* – intuitions which find their expression in the well-known words of Aristotle's *Metaphysics*

To say what is that it is, or what is not that it is, is false while to say of what it is that it is, or what is not that it is not, is true.

If we wished to adapt ourselves to modern philosophical terminology, we could perhaps express this conception by means of the familiar formula:

The truth of a sentence consists in the agreement with (or correspondence to) reality.

(For a theory of truth […] upon the latter formulation the term "correspondence theory" has been suggested).

If, on the other hand, we should decide to extend the popular usage of the term "*designate*" by applying it not only to names, but also to sentences, and if we agreed to speak of the designates of sentences as "states of affairs," we could possibly use for the same purpose the following phrase:

A sentence is true if it designates an existing state of affairs.

However, all these formulations can lead to various misunderstandings (or none of them is sufficiently precise and clear (though this applies much less to the original Aristotelian formulation than to either of the others); at any rate, none of them can be considered a satisfactory definition of truth. […].

(c) Our understanding of the notion of truth seems to agree essentially with various explanations of this notion which have been given in philosophical literature. What may be the earlier explanation can be found in Aristotle's *Metaphysics*:

To say of what is that it is not, or of what is not that it is, is false while to say of what is that it is, or of what is not that it is not, is true.

The intuitive content of the Aristotelian formulation appears to be rather clear. Nevertheless, the formulation leaves much to be desired from the point of view of precision and formal correctness. For one thing, it is not general enough: it refers only to sentences which «say» about something «that it is» or «that it is not»; in most cases it would hardly be possible to cast a sentence in this mood without slanting the sense of the sentence and forcing the style of language. This is perhaps one of the reasons that in modern philosophy various substitutes for the Aristotelian formulation have been offered. As examples we quote the following:

A sentence is true if it designates an existing state of affairs.

The truth of a sentence consists in its conformity with (or correspondence to) reality.

Due to the use of technical philosophical terms these formulations have undoubtedly a very «scholarly» sound. It is my feeling, however, that the new formulations, when analyzed more closely, prove to be less clear and unequivocal than the one given by Aristotle.

The conception of truth which has found its expression in the Aristotelian formulations (and in related formulations of more recent origin) is usually referred to as the *classical*, or *semantic conception of truth*; the semantic character of the term «true» is clearly revealed by the explanation offered by Aristotle and by some formulations which will be given in our further discussions. One speaks sometimes of the correspondence theory of truth as the theory based upon the classical conception.

> We shall try to obtain here a more precise explanation of the classical conception of truth, which could supersede the Aristotelian formulation preserving its basic intentions.

Although Tarski was explicit about his philosophical affiliations with Aristotle, not all agree that **STT** follows the tradition originated by the Stagirite (or any other traditional account of truth). In the Introduction I mentioned the critical comments of Black and Putnam about the philosophical content of Tarski's construction. In Hilbert, Bernays 1939, p. 278, we can read that is it not the case that if someone employs the name 'the definition of truth' (*Wahrheitsdefinition*) for a given explanation of what truth is, the explication offered, provides a philosophical clarification (*philosophische Aufklärung*) of this concept; it is rather unquestionable that this characterization is meant to allude to Tarski as well (Tarski 1933 is quoted a few lines this point is made). Roman Ingarden (see Ingarden 1949, p. 304), although he considers Tarski's work on truth as interesting and important, maintains that the explanation of the concept of truth in Tarski 1933 is limited to a few "curtly platitudes". To quote Putnam 1985–1986, p. 333:

> As a philosophical account of truth, Tarski's theory fails as badly as it is possible for an account to fail.

Such opinions, expressed by top logicians or philosophers, cannot be ignored and require a closer analysis of how **STT** is successful as a philosophical enterprise. Tarski himself, perhaps somehow annoyed by such assessments, wrote (Tarski 1944, p. 361)

> In general, I do not believe that there is such a thing as "the philosophical problem of truth". I do believe that there are various intelligible and interesting (but not necessarily philosophical) problems concerning the notion of truth, but I also believe that they can be exactly formulated and possibly solved only on the basis of a precise conception of this notion.

On the other hand, Kokoszyńska in her letter to Tarski on June 17, 1950 (the original can be found in the Bancroft Library in Berkeley) said "I sometimes doubt whether you properly appreciate the philosophical consequences of your theory of truth"; it is interesting that this letter was written five years after the publication of Tarski 1944. Kokoszyńska was right because there is much more to say about the philosophical consequences of **STT** than one can find in Tarski's writings.

The views of Tarski expressed in the quotations (a)–(c), are similar but still differ in detail. Three points of agreement are the following: (i) the reference to Aristotle's intuitions as the most proper starting point for seeking out the required truth-definition; (ii) considering Aristotle's formulation as semantic (although (b) lacks this factor, but we can assume that it is tacitly present); (iii) the claim that the intuitions expressed by the Stagirite require several improvements. As concerns differences, Tarski characterized the classical theory as identical with the correspondence theory in (a), but he was more careful in (b) and (c). In 1933, he considered the formulation 'a true sentence is one which says that the states of affairs is so and so and the state of affairs indeed is so and so' as a proper rendering of Aristotelian semantic intuitions, but the similar formulations mentioned in

(b) and (c) are put forth with some reservations namely that, although "Due to the use of technical philosophical terms these formulations have undoubtedly a very «scholarly» sound", they are less precise than Aristotle's explanations in *Metaphysics*. Tarski informs us (Tarski 1933, p. 155, footnote 2) that "a very similar formulation can be found in Kotarbiński". Let us take it for granted that the sentence used by Tarski is very similar to (Db) in Chap. 3, Sect. 3.7. However, the Polish logicians and philosophers mentioned in Chap. 3, Sect. 3.9 did not accept with identifying of the classical conception of truth with the correspondence theory of this concept. Thus, it seems that Tarski resorted, the "folk" philosophy, so to speak, when he proposed that identification in question. The change of his view may have been be caused by systematic difficulties in the understanding of what appears to have s "a very «scholarly» sound." It is also possible that he was influenced by Kotarbiński 1934 and discussions with Kokoszyńska (she is mentioned in Tarski 1944, p. 665), who explicitly argued that **STT** rehabilitate the classical (Aristotelian) truth-definition.

I shall consider **STT** as a version of the classical truth-theory (**CTT** for brevity). The question that now arises is which element of **STT** can be regarded as a bridge between the semantic conception of truth and the Aristotelian tradition. Tarski argued repeatedly that the main intuitive idea of **CTT** is covered by **T**-sentences. Thus, to invoke his paradigmatic example, the sentence "snow is white" is true because it says that snow is white, and snow is white. And this content has its proper rendering in the corresponding **T**-sentence.

(1) the sentence 'snow is white' is true if and only if snow is white.

How to relate (**Df2**) or **(Df3)** (the official **SDT**; see Chap. 8) to the intuition expressed by (**TS**) (it is enough to use this simplified version of **T**-scheme)? The answer is offered by **CT** (see Chap. 7, Sect. 7.5), which requires that every instantiation of **(**T**S**) logically follow from **SDT**. Although (**TS**) cannot be considered as a truth-definition, its particular forms, like (1), might be regarded as partial definitions; for instance, (1) provides a partial truth-definition for the sentence 'snow is white'. Continuing this line of reasoning, one could say, the conclusion about the status of the official **SDT** would assert that if the consequences of a definition agree with intuitions, the same applies the definition itself, at least to some extent. This last reservation is justifies because if *A* logically entails B, the content of the consequent can be smaller than the content of the antecedent.

Due to the last statement, the above argument does not explain directly the status of **SDT** as such, that is, saying that a sentence is true given that it is satisfied by all infinite sequences of objects, at least one such sequence or the empty sequence. In fact, Tarski never used his official **SDT** as carrying intuitions concerning the concept of truth. This fact suggests that he understood satisfaction by all infinite sequences, by at least one such sequence, or the empty sequence as various technical constructions devoid of any intuitive and philosophical content, although Tarski he made that assertion with respect to the role of the empty sequence. Can we look for ordinary intuitions associated with the official **SDT**? My claim is that we can. Extending the above remark about the connection between the content of the antecedent and consequent in the case of the entailment-relation, intuitive consequences can be logically entailed by assumptions which are intuitively neutral, or even counterintuitive for the first sight, nothing precludes that information covered by the consequences of some premises, is somehow, even tacitly, present in premises themselves. If so, we can expect that a proposed definition **D** of a concept **C** should exhibit at least some of the intuitions captured by consequences of **D**. In our case, that means that the intuitive (philosophical) content of **T**-sentences—namely that a sentence saying so and so is true if and only if the sentence in question says so and so, and this it is so and so—has a counterpart in the satisfaction of this sentence by all infinite sequences of objects, etc.

As I noted earlier (see **DG4VIII**), it is preferable to use satisfaction by all infinite sequences of objects as the definiens in **SDT**. Independently of whether it was a casual circumstance or something intuitively motivated that Tarski himself choose this conceptual path, I shall adopt this route as basic, also because it was followed by some commentators. Assume that we look for an explanation of correspondence or adequacy to (with) reality, and interpret the infinite sequences of objects as facts (I remember such attempts, but, I am unable to name specific authors). Even if one attempts to consider sequences as facts, this way of explaining the nature of truth seems completely misleading. First of all, even given possibility that we can always truncate infinite sequences to finite subsequences that correspond to free variables in formulas, sequences are nothing more than elements of relations. That means that if s is an n-termed sequence, it belongs to some n-ary relation, and it is altogether unclear what that has to do with facts in their ontological meaning, that is, as the pieces of reality. Secondly, even if we ignore the last remark, then, due to the equivalence of satisfaction by all infinite sequences and by at least one sequence, we have a the consequence that if A is true, it is satisfied by a fact if and only if it is satisfied by all facts. Thirdly, and even worse, due to the equivalence of satisfaction by all sequences (at least one sequence) with satisfaction by the empty sequence, A is true if it is satisfied by the empty fact, provided that sequences are interpreted as facts. Yet the idea of the empty fact looks extremely curious. Fourthly, since sentences have no free variables, if we intend to avoid satisfaction by all sequences, we have to address the problem of which sequences correspond to true sentences. However, this step makes the concept of being false very problematic, since the intuition that false sentences are not satisfied by any sequence is stable and looks reasonable. Fifthly, the account of truth as satisfaction by all sequences of objects (which is required for the equivalence of various versions of the official **SDT**) immediately implies the slingshot argument (see Chap. 3, Sect. 3.6) that all true sentences are satisfied by the same facts and, ultimately, by the Great (Big) Fact, that is, the Reality (see Neale 2001, pp. 49–57). In fact, such a critique is directed against many truth-theories based on the concept of correspondence with facts. Since I will not use this approach to truth, I feel justified in ignoring the slingshot argument in my further considerations (in particular, I shall not comment on the defence of facts in Neale 2001). It is enough to observe that since sequences are not facts, **STT** does not falls under this objection.

My interpretation of **STT** adheres the distinction of the weak and strong interpretation of the correspondence relation (see Chap. 3, Sect. 3.7.8; see also Niiniluoto 1994 for a discussion of objections against interpreting **SDT** as a correspondence theory and Niiniluoto 1999 for comparisons of Tarski on truth with logical empiricism). The latter explains this relation via a structural similarity (isomorphism, homomorphism, copying, picturing, etc.) of facts to bearers of truth; the views of Husserl, Russell and Wittgenstein view (see Chap. 3, Sect. 3.8) are examples of the strong interpretation. Tarski rejected such an approach. My hunch is that his reservations against "a very «scholarly» sound" involved the conviction that formulations a la 'designation of a state of affairs' are too strong from the ontological point of view and do not properly reflect the phrase 'things are so and so'. However, this conjecture as negative does not propose any positive understanding of correspondence in the weak sense. As I remarked in (**DG13VIII)**, I neglect attempts via references (designations, denotations) of sentences in models, because they are more or less artificial additions to model theory.

I proposed at the end of Chap. 7, Sect. 7.6 to understand the provisional truth definition ((20) in Chap. 7) as saying how things are (this expression means the same as 'things are so and so'). To develop this approach let me start with the concept of satisfaction. This path is reasonable since for sentence truth is a generalization (or modification, if one prefers) of satisfaction. Consider the simplest formula, that is, $P(x)$ with x as a free variable. Assume that the interpretation of predicates is fixed; in addition, we exclude tautologous and contradictory predicates. Let (i) 'x is a city' serve as an illustration. If we substitute x by 'London', we obtain (ii) 'London is a city', a true sentence, due to the fact that $\mathbf{v}(x)$ = **London** and this object is a city. If $\mathbf{v}(x)$ = **Thames**, we obtain (iii) 'Thames is a city'—which is false. We say that (ii) says how things are, but (iii)—not, and that these semantic conventions depend on how x is valued. Whereas the preceding is a semi-technical jargon stimulated by logic and semantics, illustrations (ii) and (iii) come from ordinary language. Omitting special situations like ambiguities, indexicals, polemics, etc., fixing values of predicates and individual constants are directly determined by meanings of expressions, and are taken for granted. In particular, variables are practically not used except of special occasions, for instance in the teaching of mathematics. The practice of our colloquial language causes that when we utter (ii), we do not think that $\mathbf{v}(x)$ = **London**. Even more, we normally do not observe that 'London' is valued by **London**. Hence, the difference between satisfaction and truth is vague in the use of ordinary language. The question 'Does the form (condition) 'x is a city' become true after substituting the proper name 'London?' for the variable x looks as artificial—apart from its function in elementary logic classes.

The above analysis of (i)–(iii) does not suffice in formal semantics. Nevertheless, ordinary illustrations can be extended to truth in a model **M**. In order to do so, one needs to go deeper into the analysis of (i). Let x_1 be the first variable in the list of all variables of $\mathbf{L_{FOL}}$; the cardinality of $\mathbf{U^M}$ = the cardinality of the set of individual variables = $\boldsymbol{\aleph_0}$. An object $\boldsymbol{s_1}$ from $\mathbf{U^M}$ either satisfies the formula $P(x_1)$, or it does not. Since satisfaction (or not) of this formula only depends of how the variable x_1

is interpreted (recall that we assume that the interpretation P is fixed), the one-term sequence, $<\boldsymbol{s_1}>$, of objects can be extended by adding the unlimited (in fact, infinite but denumerable) number of new terms. Such extensions do not influence the valuation of x_1; the interpretation can be changed—but only by correlating a new object from $\mathbf{U^M}$ with this variable. Extending $<\boldsymbol{s_1}>$, in so far as we are working with $P(x_1)$, is only a technical device treating uniformly treating of formulas of arbitrary length. Anyway, the behaviour of formula $P(x_1)$ is not stable with respect to elements of $\mathbf{U^M}$, because this formula is at times satisfied, at others not. So we cannot say how things are in **M**, because in some circumstances they fall under the predicate P, but in others they do not. The elimination of free variables in an arbitrary formula A, if finished, completes the semantic stabilization of semantic properties consisting in the possession of a definite logical value (not necessary—the truth) by A.

However, one might ask why satisfaction by all infinite sequences of objects (or its equivalents) is to be identified with truth, as well as why satisfaction by no such sequence is to be considered as defining falsehood. Moreover, if A is a sentence, it is satisfied by all sequences or by no sequence (see Chap. 8(8)). In all the preceding explanations of the concept of truth, I proceeded from satisfaction to truth. Now I shall employ the reverse path and try to show how an understanding of truth contributes to the technical understanding of satisfaction. Assume we explain that '$P(a)$' is true by saying that formula $P(x_1)$ is satisfied by the object **a** due to the fact that $\mathbf{v}(x_1) = \mathbf{a}$. According to **SDF**, $P(a)$ is made (incidentally, this formulation should not be interpreted via the idea of so-called truth-makers'; see remarks at the end of this section) true by all infinite sequences of objects or, due to Chap. 8(6)), by at least one such sequence (I ignore at this point the empty sequence). Here is the list of these sequences:

(*) $\boldsymbol{s_1}, \boldsymbol{s_2}, \ldots, \boldsymbol{s_{k-1}}, \boldsymbol{s_k}, \ldots$
$\boldsymbol{s_2}, \boldsymbol{s_1}, \ldots, \boldsymbol{s_{k-1}}, \boldsymbol{s_k}, \ldots$
$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$
$\boldsymbol{s_k}, \boldsymbol{s_{k-1}}, \ldots, \boldsymbol{s_2}, \boldsymbol{s_1}, \ldots$
$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

If we say that the truth of a sentence is just independent of the valuation of free variables, it does not mean that they (even bound variables) are not valued by objects from $\mathbf{U^M}$, or they lost the initial interpretation. Since the language **L** is interpreted, all components of its alphabet have a semantic interpretation. Sequences from the list (*) provide all possible valuations of variables. For the formula $P(x_1)$ the valuation of x_1 is relevant. Although it must be equal to $\boldsymbol{s_1}$ in every sequence, different individuals can be selected as instantiations of $\boldsymbol{s_1}$. Assuming '$P(a)$' is true, we obtain $\mathbf{v}(x_1) = \mathbf{v}(a) = \mathbf{a}$, for any infinite sequence (or at least one sequence). Consequently, (*) can be rewritten as

(**) $\mathbf{a}, \boldsymbol{s_2}, \ldots, \boldsymbol{s_{k-1}}, \boldsymbol{s_k}, \ldots$
$\boldsymbol{s_2}, \mathbf{a}, \ldots, \boldsymbol{s_{k-1}}, \boldsymbol{s_k}, \ldots$
$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

$\boldsymbol{s_k}$, $\boldsymbol{s_{k-1}}$, …, $\boldsymbol{s_2}$, **a**, …
……………………

Now, if $\mathbf{v}(a) \in \mathbf{P}$, P(a) is true. What happens, when $P(a)$ is false in the above sense? That would mean that $\mathbf{v}(a) \neq \mathbf{a}$ for every sequence in the list (**). However, it also means that the formula $P(x_1)$ could not be satisfied, even though it was assumed to generate '$P(a)$' by a the substitution of 'x_1' by 'a'. This is an argument for a non-trivial semantic import of satisfaction by all infinite sequences. Looking at (*) and (**), we see that all sequences are mutual permutations and none of them is distinguished or privileged a priori. Thus, if we take any sequence from the list (**) as an illustration, we can employ any other to the same effect. This observation provides a reason that satisfaction by at least one sequence is equivalent to satisfaction by every sequence. As far as the empty sequence is concerned, if admitted at all, it is a subsequence of any other sequence (which follows from (**A1**)–(**A4**) in Chap. 8, Sect. 8.2). If this is the case, the sequence in question can be extended by adding an infinite number of terms. If A is a sentence, its satisfaction by at least one of these extensions is necessary, but if by one, then the same applies to all. Further, if a sequence <**s**> satisfies A, no subsequence of <**s**> can make A false, because if it did no sequence would satisfy this sentence, contrary to the assumption. Returning to (i)–(ii) as concrete illustrations, the sentence 'London is a city' is true, due to (a) **v**(London) = **London** and (b) **London** $\in$ **City**. On the other hand, the formula 'x_1 is a city' is satisfied by **London** (and many other objects too), but in order to be satisfied by **London**, the equality $\boldsymbol{s_1}$ = **London** must hold for at least one sequence from the list (*).

(**DG1**) The above argument can first be generalized to atomic formulas of arbitrary arity. If we consider the formulas $P(a_1, \ldots, x_k)$ and $P(a_1, \ldots, a_k)$ the diagonals in (*) (concerning $\boldsymbol{s_1}$) and (**) (concerning **a**) become so-called k-dimensional diagonals. The next generalization involves complex formulas. Things become simpler in canonical models and Henkin models, because every object in **U** has its own individual name.►

Sequences taken in isolation from sentences do not determine logical values. This circumstance is clearly indicated by saying that if $\mathbf{v}(a) \in \mathbf{P}$, P(a) is true. The construction of semantics for **FOL** in Chap. 5, Sect. 5.2.4 explains what is going on, because the valuation of the alphabet **AL** has no direct extension to propositional formulas (open and closed). In particular, extensions of predicates are not given by sequences of objects used in **SDT**. For this reason, the complete description of satisfaction and truth as semantic properties requires using the following statement: 'a formula A is true (satisfied) in a model **M** under interpretation $\mathfrak{I}$'. If we keep that in mind, the official **SDT** is not merely a formal trick (to some extent it is), but also has a quite reasonable intuitive content. Thus, we can say that A, if true according to **SDT**, says how things are in **M** or how they are not, and, if A is false, it says nothing about these matters. Since **SDT** entails **T**-equivalences, we can now say that their intuitive plausibility also has its source in the truth-definition itself, which shows not only that a given formula is semantically stable (see above), but also when semantic stability becomes truth or falsehood. In particular, we do not need to invoke traditional explanations of what the essence of the correspondence relation in its strong sense is. Thus, **STT** can be qualified as naturally based on the weak sense of the relation that holds between truth and reality. This relation can also be termed 'semantic correspondence'. Whereas this label is used in Hill 2002 as having the deflationary meaning, my usage is more substantive (see the remarks on minimalism in Sect. 9.7 below). To complete this discussion, Tarski never said that **SDT** is the only truth-definition that satisfies **CT**. His view was that any materially correct truth-definition should entail **T**-equivalences for all sentences of **L**. For instance, the treatment of states of affairs proposed in Chap. 8, Sect. 8.7 also leads to the material adequacy of the related truth-definition, presumably involving a version of **CTT**, but I shall not pursue this issue. At any rate, **STT** is classical, substantive, non-epistemic (see Chap. 4, Sect. 4.4) and relational in Russell's sense (see Chap. 3, Sect. 3.3).

(**DG2**) The fact that the valuation of **AL** cannot be recursively extended to valuation of open formulas and sentences has a certain unexpected philosophical import, because it might be associated with presenting and judging as different mental acts. If someone says that he or she has a presentation of such and such object, one is reporting on a reference to an object or its properties. On the other hand, judging involves that a given person presupposes that the judgement in question is true or false. According to Brentano and his students (also Twardowski), judgments (propositions) are not combinations of presentations (the allogenic theory of judgements), but mentalities (mental facts) *sui generis* (the idiogenic theory of judgements). Semantics for **FOL** displays well the latter theory of judgements (propositions, sentences) by defining satisfaction (truth) recursively starting with clauses for atomic formulas, although interpretation of variables, individual constants and predicates is assumed as given in advance. That means that interpretation itself without conditions of satisfaction does not suffice for generating logical values of sentences. Psychological speech is employed even in advanced textbooks of mathematical logic. For instance (see Goldstern, Judah 1995, p. 57), **ML** is considered as the language of a subject.►

(**DG3**) Here is (incomplete to be honest) list of books dealing with **CTT** (and other contemporary truth-theories) and also discussing some problems related to **STT** as a version of the correspondence (or classical) theory: Armour 1969, O'Connor 1975, Williams 1976, Moreno 1992, Kirhham 1993, David 1994, Twardowski, Woleński 1994, Alston 1996, Soames 1999, McGrath 2000, Weingartner 2000, Hill 2002, Newman 2002, Künne 2003, Vision 2004, Burgess, Burgess 2011, Frápolli 2013, Pedersen, Wright 2013, Rasmussen 2014, Achourioti, Galinon, Fernández, Fujimoto 2015; these books discuss many other questions associated with STT, for instance, truth and meaning, truth and ordinary language, etc.►

As I noted in Chap. 4, Sect. 4.5, the problem of a truth-criterion (**CrT**, for brevity) is considered serious for every **CTT**. As far **SDT** is concerned, it does not imply any specific truth-criterion. Tarski himself (see Tarski 1969 and Chap. 8, Sect. 8.5) considered proof as the truth-criterion in mathematics, but, due to (**TT**) as only approximate. However, the central role of proof in mathematics is independent of **SDT**, and it is difficult to see Tarski's view as solving the general problem of **CrT** in the context of **STT** even in the case of mathematics. But if we pass to empirical statements like (ii), the method of (deductive) proof is not sufficient or usable in most cases. From my perspective, I propose a solution suggested by Anna Kanik, my former student (unfortunately, she resigned from doing philosophy). Assume that we attempt to check the truth of a sentence of the form $P(a)$. We replace it by $P(x)$ and ask how to check that $\mathbf{v}(x) = \mathbf{a}$. In other words, every method which allows us to answer that $P(a)$ is true, because the object **a** satisfies the formula $P(x)$, provides **CrT** in a given situation. Such a method can consist of proof via deduction, perceiving, remembering, using analogy, comparing by coherence, relying on utility or evidence, or appealing to consensus. **STT** does not require only one criterion, or a claim that **CrT** is ultimate. Perhaps an interesting feature of this approach consists in accommodating to **STT** various so-called non-classical truth-theories as offering different complementary criteria of truth. It still remains here to respond to objections raised by sceptics. Roughly speaking, the sceptic says that no **CrT** is impossible to use it. Assume that A is a sentence to be checked as true or false by a given **CrT**. Hence, since the use of **CrT** requires the statement '**CrT** is suitable for being the truth-checker', we fall into circularity. However (see Ajdukiewicz 1949), a closer analysis shows that this premise is not necessary, because to use a device for achieving a cognitive (and any other) task does not depend of the knowledge that the device employed is suitable in the given situation. The sceptics strengthened their view by claiming that the proper **CrT** should be ultimate and infallible. However, this claim was inherently associated with the sceptical attack against the idea of *episteme* in Parmenides' sense. In conclusion, the situation of **STT**, as far as the issue of **CrT** is concerned, is at least no worse than that of other truth-theories—in particular, of correspondence theories.

Brentano (see Chap. 3, Sect. 3.5) formulated several arguments against **CTT**, namely:

(a) The correspondence definition of truth does not meet the objections raised by sceptics and relativists;
(b) The correspondence theory of truth cannot explain what corresponds to negative judgements, in particular, to negative existential ones;
(c) The correspondence theory of truth does not explain why theorems of logic and mathematics are true, because such statements do not correspond to specific objects, but are universally valid;
(d) The correspondence theory of truth cannot meet the objections of circularity, infinite regress or petitio principii;
(e) The notion of correspondence is too vague in order to constitute a satisfactory foundation of a truth-theory.

Objections (a) (more precisely, its part related to scepticism; relativism will be considered later), (b) (to repeat, negative statements, if true, are true in **M** in the same sense as positive ones), (c) (to repeat, logical truths are true in all models and this account is entirely consistent with **SDT**) and (e) (to repeat, the concept of strong correspondence is actually too vague, but **SDT** does not use it, because it employs the concept of weak correspondence) were already addressed in earlier fragments of this book (see also Woleński 1989a).

Thus, it remains to account for (d), although it was too partially answered (see Chap. 7, Sect. 7.6) by pointing out that **SDT** does not employ the resources of **L** or, more precisely, that this truth-definition is based essentially on specific concepts of **ML**. The objections of infinite regress and petitio principii can be dealt with together. They point out that either **SDT** must be earlier applied for **ML** in advance, although this step leads to infinite regress, or if **SDT** is not available for **ML**, the entire procedure falls into *petition principia.* However, it is actually not true that in order to construct **SDT** for **L** one must have a similar definition for **ML**. Yet one may claim that adopt **STT** one needs to assume that certain theorems from set theory and logic are true. Although that is true, this situation appears to be customary in science. For example, if we argue that quantum mechanics explains several facts that obtain in the macro-world, we assume that all measurements grounding this theory are realized by instruments based on classical physics. In general, this example shows that our cognitive practices, if they are described rigorously, require an intuitive understanding of something else. In the case of metamathematics, investigations on a formalized **L** are performed in a partially informal **ML**. There is no way out of this situation. Clearly, some philosophers claim—as Husserl did—that philosophical investigations should be free of all presuppositions (in this case, the analysis of **L** in **ML** cannot employ logical tools defined in the former), but this project is more utopian than effectively realizable in cognitive activities.

Now I would like to comment briefly comment the theory of truth-makers, sometimes regarded as a modification of **STT**—but sometimes as an alternative to it. The initial idea is such (see Mulligan, Simons, Smith 1984, p. 10/11):

> […] an adequate account of truth must include considerations which are other than purely semantic in the normally accepted sense. Our suggestion here – a suggestion which is formulated in a realist spirit – is that the way to such a theory lies through direct examination of the link between truth-bearers, the material of logic, and truth-makers, that in the world in virtue of which sentences or propositions are true.

More specifically, the theory of truth-makers is based (see Rami 2009a, p. 3) on the so-called truth-maker principle in the following form

(**TMP**) For every x, x is true if and only if there is a y such that y is a truth maker for x.

This principle does not appeal to models in the semantic sense, but directly to the real world. Sometimes it is supplemented by the claim that if a sentence A is made true by a truth-maker the relation of truth-making is necessary. I see two problems for the account of truth via the concept of truth-makers. The first question is as follows. If truth-makers are to be components of the real world (facts, for instance), we should understand them mereologically (according to Leśniewski's mereology, that is, the theory of parts and wholes) rather than set-theoretically. If this suggestion is adopted, the problem arises of how large (what is their scope?) are truth-makers. Assume that we consider the sentence (a) 'Jan Woleński lives in Poland'. This sentence is made true by several facts, for instance, **Jan Woleński lives in Cracow**, **Jan Woleński lives in Małopolska** (an administrative province in which the city of Cracow is located), etc. Which segment of reality constitutes the truth-maker for (a)—the minimal, or some other? If the minimal, which is that? Secondly, the idea truth of sentences as necessitated by truth-makers appears very unclear. Is it logical necessity or real necessity? These difficulties need to be overcome in order to achieve a basis for defining truth via truth-makers. I see more promise in viewing truth-makers as associated with the problem of truth criteria. Although my discussion of (**TPM**) and related problems is limited, and thereby simplified (see Armstrong 2004 for a more extensive discussion), it suffices for concluding that the fusion of the theory of truth-makers and **STT** is very problematic.

(**DG4**) **CTT** is commonly considered as the most intuitive from the ordinary point of view. Arne Naess (see Naess 1938, Naess 1953) investigated empirically how educated Norwegians understood truth. He asked about the main philosophical truth-theories (correspondence, pragmatic, etc.). The result was that all of them achieved comparable endorsement, but none exceeded 20%. In fact, **CTT** gained the highest popularity, but not sufficient to be granted absolute prevalence as the most intuitive understanding of truth. Naess also asked about the relation between (a) the sentence *A*, and (b) the sentence '*A* is true'. More than 90% of respondents answered that (a) and (b) are just synonymous. Moreover, more than 50% acknowledged that the negations of (a) and (b) are synonymous. Incidentally, the differences between both results show that the understanding of negation and denials is encumbered with some additional problems. Naess (Naess 1953, p. 41) remarked that statistical data confirm the hypothesis that the semantic theory of truth agrees with ordinary linguistic intuitions (see also Tarski 1944, p. 684; see Ulatowski 2016 and Barnard, Ulatowski 2016, Ulatowski 2017 for various and more detailed comments on Tarski on Naess). Another research was carried out in the USA in 2009 (see Scaglia 2011, p. 24). The entire group of questioned respondents consisted in 3226 persons, including 1803 members of faculties of philosophy and 829 graduate students in philosophy. The result was that "44% accepted or lean towards correspondence theories, 20.7% accept or lean towards deflationary theories and 13.8% epistemic theories". This result is not quite transparent, because we do not know of how the questions were formulated and how being a professional philosopher or a graduate philosophy student determined responses. At any rate, the dominance of **CTT** is unquestionable in the research in 2009.►

9.3 The Status of T-Equivalences

By the problem of the status of **T**-sentences I understand the question as to whether they are analytic, tautological or material equivalences. The answers given to this question by various authors differ fundamentally. Although Tarski himself did not consider this problem, many authors try to give interpretations of his possible view that I discuss in this section. For instance, many commentators argue that **T**-equivalences are analytic or even tautologous. On the other hand, Tarski's well-known skepticism about the analytic/synthetic distinction suggests that we should be very careful about attributing to him the view that **T**-sentences are analytic. Quine (see Quine 1953a, p. 137, note 9) remarks that Tarski "has not claimed that [**T**-sentences] are analytic" and that this fact is "sometimes overlooked". However, this is only a negative statement and does not contribute toward a positive perspective. In particular, Quine's opinion does not entail that **T**-equivalences are synthetic, because he rejected the existence of pure synthetic or analytic statements. Michael Dummett's view is more definite on this point. He says (see Dummett 1978, p. XX) that for Tarski, every **T**-biconditional is "no more than material equivalence, i.e. identity of truth-value". Van McGee (see McGee 1991, p. 1) maintains that one could be inclined to regard **T**-equivalences as "not only true, but analytic". Unfortunately, McGee continues, this very simple and intuitive thesis is ruled out by the Liar Paradox. For Scott Soames (see Soames 1995) **T**-equivalences are neither analytic nor a priori. On the other hand, if analytic sentences are defined as asserted by any competent user of a given language, the extended (**TS**), that is, the equivalence ⌜'*A*' is true in my language if and only if *A*⌝, is analytic. The same author (see Soames 1999, p. 106) says that statements of the form 'a proposition *A* is true if and only if *A*' are necessary, a priori and trivial, due to an analytic content-connection between their parts; Soames informed me in a letter dated on October 14, 2002) that he was never entirely convinced of this view about **T**-sentences and that he rejects it now.

Putnam (see Putnam 1985–1986, pp. 332–333) argues that (**TS**) belongs to the theorems of logic valid in **ML**, and thereby every one of its instances presents a truth valid in all possible worlds. Such a way of treating **T**-equivalences has its justification in the fact that they are true by virtue of logical axioms and axiomatically assumed conventions for forming names of sentences. Marian David (David 1994, pp. 130–135, David 2008) regards the instances of (**TS**) as contingent (not necessary), because its components, that is, **T**(*A*) and *A* are contingent. Consider Tarski's favorite example, that, is (1). Assume that this sentence is necessary. However, the sentence 'snow is white' might mean that grass is green. This would entail the falsity of (**TS**) vis-à-vis the assumption. Consequently, (**TS**) itself cannot be necessary. Volker Halbach (Halbach 2011, Chap. 3) endorses the same view pointing out that **T**-equivalences are contingent "as they depend on what sentences express or how sentences are used". Yet David remarks that a proper truth-definition should be "something more" (in general, a conventional element; thus, (**TS**) is a conceptual truth by convention) than a contingent assertion, and the same applies to concrete **T**-biconditionals as its logical consequences. Hence, (**TS**) cannot function as a good truth-definition; this conclusion is intended as critical against minimalism (see also Sect. 9.7 below). On the other hand, Halbach says that if **T**-sentences are adopted as axioms as the case is in deflationism, they "need to be necessary".

Künne (see Künne 2002, p. 183) argues that Tarski should (not that he did) regard **T**-biconditionals as necessary truths. Künne's main claim runs as follows (p. 183):

> Criterion [= Convention – J. W.] T demands that the pertinent T-equivalences follow from the definition of 'true' for L (plus some non-contingent syntactical truths like "snow is white" is not identical with "blood is red"). The definition itself is not a contingent truth, for it is constitutive of L that its sentences mean what they do mean. Now a conceptual truth *A* cannot entail a contingent truth *B*; for otherwise, by contraposition, the negation of *B*, which is as contingent as *B*, would entail the negation of *A*, which is itself conceptually false. […]. Hence, Tarski ought to accept [that (**TS**)] is necessary.

Let us denote a conceptual falsity by **CF**. The principle expressed in the final fragment of the quotation from Künne has the following symbolic formulation:

(2) If $B \in \mathbf{CF}$ and $A \vdash B$, then $A \in \mathbf{CF}$.

Denote a conceptual (analytical, necessary, tautological, etc.—these qualifications are merely provisional) truth by **CTr**. Assuming bivalence, we have:

(3) If $A \in \mathbf{CTr}$ and $A \vdash B$, then $B \in \mathbf{CTr}$.

Although the relation between (2) and (3) does not raise any doubt, the final conclusion (asserting the necessity of (**TS**)) depends on the concept of conceptual truth and its relation to the notion of tautology. The problem appears serious, because if (**TS**) and its consequences are tautological, **T**-equivalences cannot be regarded as bearing any intuitive content.

If **T**-equivalences belong to **ML**, their status must be investigated at the metalinguistic level. The foregoing discussion suggests two extreme standpoints concerning the status of instances of (**TS**): (a) **T**-biconditionals are merely material equivalences (Dummett); (b) **T**-biconditionals are logical tautologies (Putnam). I will show that these two positions cannot be accepted (see Woleński 2001 for a critique of Putnam; see also Woleński 2012, Woleński 2014). For the sake of argument, let us suppose that a concrete **T**-sentence—for example, (1)—is just a material equivalence. In addition, we assume that the sentence 'snow is white' is true. If both provisos are satisfied, we can replace 'snow is white' on the right side of (1) by another true sentence, for instance, 'blood is red'. This produces the next true equivalence:

(4) The sentence 'snow is white' is true if and only if blood is red.

It is frequently maintained that **T**-sentences express or establish, at least under their standard interpretation, truth-conditions for sentences. Hence, (1) states the truth-condition for 'snow is white', but (4) does not. Although (1) and (4) are material equivalences, according to the semantic theory of truth, they should be equivalent with respect to their provability in **ML**. This requires demonstrating that (4) is provable in **ML**. In order to do that, we can assume that the expression 'snow is white' just means that blood is red. Nothing prevents us from adopting this interpretation, but this step changes essentially our intuitive and tacitly assumed semantics. If 'snow is white' means that snow is white, the new reading of the sentence in question changes the original semantic equipment of language into such in which 'snow is white' means that blood is red. More importantly, this would change a language **L** for which our truth-definition is elaborated into a language **L′** with a different interpretation function. Three points seems relevant here. Firstly, syntactic conventions have no a great significance here. Secondly, the language **L** for which we formulate a truth-definition is semantically interpreted. Thirdly, **T**-sentences function as theorems of **ML**. Although that three points were strongly stressed by Tarski, the second is very frequently overlooked and commentators usually consider the semantic definition of truth as directed exclusively at formal languages. In fact, one should distinguish between formal and formalized languages.

If the expression $\mathbf{T}(A)$ is rendered 'it is true that A', the letter **T** now denotes a monadic sentential connective forming a new sentence with a referent of A as its argument (see Chap. 5(4b)). The formula $\mathbf{T}(A) \Leftrightarrow A)$ is a demonstrably logical propositional tautology on this reading, and it was regarded as such in the algebra of logic. For example, Louis Couturat (see Couturat 1905, p. 84) adopts the principle of assertion (in his terminology) $(a = \mathbf{1}) = a$ (**1** refers to the constant true sentence) and adds the following comment:

> To say that a proposition a is true is to assert the proposition itself. In other words, to state a proposition is to affirm the truth of that proposition.

In the symbolism I employ, the principle of assertion—regarded by Couturat as a peculiar or characteristic formula of the algebra of propositions—can be written as $(A \Leftrightarrow \mathbf{1}) \Leftrightarrow A$. Of course, this move requires a slight extension of the typical language of propositional calculus by adding the propositional constant **1** as a new connective. Moreover, the definition of a well-formed formula must be suitably modified. However, this treatment has a restricted significance and does not satisfy Putnam's tasks. Since his thesis qualifying **T**-equivalences as tautologies (formulas valid in all possible worlds) is not restricted to syntactic items of propositional logic, (b) should be justified or denied with respect to the form 'A is true'. At the first glance, this structure can be converted to 'it is true that A'. However, if **T** is a connective, the formula $\mathbf{T}(A) \in \mathbf{L}$ ($\mathbf{T}(A) \notin \mathbf{ML}$). Yet Putnam explicitly states that his interest lies in **T**-equivalences as metalinguistic statements. In order to make the further reasoning more explicit, I shall employ '$A \in \mathbf{VER}$' instead of the combination $\mathbf{T}(A)$.

Assume that the formula '$A \in \mathbf{VER} \Leftrightarrow A$' is a theorem of logic. We have thus (I drop quotation marks in further considerations—recall that the phrase '$A \in \mathbf{VER}$' means 'a sentence represented by the metavariable A is true'):

(5) $\vdash (A \in \mathbf{VER} \Leftrightarrow A)$.

Let **1** refers to an arbitrary tautology; this convention does not produce any confusion with Couturat's understanding of **1**. Since Putnam claims that **T**-equivalences hold in all possible worlds, this implies that first-order logic functions as our basic logic, and, in formal semantics we can replace a problematic category of possible worlds by models. Consequently, the statement expressed by (5) is true in all semantic models or under all interpretations of its extralogical parts. Simple transformations, allowed by the principle that two tautologies are provably equivalent and the completeness theorem for first-order logic, lead directly to

(6) $\vdash (A \in \mathbf{VER} \Leftrightarrow A) \Leftrightarrow \mathbf{1})$.

Since the operation denoted by $\Leftrightarrow$ is associative, we obtain

(7) $\vdash (A \in \mathbf{VER}) \Leftrightarrow (A \Leftrightarrow \mathbf{1})$.

The distributivity of $\vdash$ over equivalence gives

(8) $\vdash (A \in \mathbf{VER}) \Leftrightarrow \ \vdash (A \Leftrightarrow \mathbf{1})$.

Observe that the property expressed by the predicate 'is a tautology' can be defined by the condition

(9) $(A \in \mathbf{TAUT}) \Leftrightarrow \ \vdash (A \Leftrightarrow \mathbf{1})$.

Assume that A is a tautology. Thus, A is universally true. If so, the right-side equivalence, that is, $(A \Leftrightarrow \mathbf{1})$, is always true, because provable. Hence, A cannot be false, which means that A is a tautology. Thereby, we can transform (5) into

(10) $\vdash (A \in \mathbf{VER}) \Leftrightarrow \ \vdash (A \in \mathbf{TAUT})$.

Although we cannot unconditionally pass from

(11) $\vdash A \Leftrightarrow \ \vdash B$,

to the formula

(12) $A \Leftrightarrow B$,

because it may happen that A is an unprovable true sentence and B is a false sentence. Since both components of (11) are false, and, in such a case, (11) is true, but (12) is false due to the normal understanding of $\Leftrightarrow$. This argument proves that passing from (11) to (12) cannot be performed in every case. Yet we can show that it is possible to omit the sign $\vdash$ under our earlier assumptions. Thus, we obtain the conclusion

(13) $(A \in \mathbf{VER}) \Leftrightarrow (A \in \mathbf{TAUT})$.

I will proceed by the *reductio ad absurdum* to justify that we should not accept (13). If we deny (13), it produces

(14) $\neg(A \in \mathbf{VER}) \wedge (A \in \mathbf{TAUT}) \vee (A \in \mathbf{VER}) \wedge \neg(A \in \mathbf{TAUT})$.

The first segment of (13), namely '$\neg(A \in \mathbf{VER}) \wedge (A \in \mathbf{TAUT})$' does not hold, because tautologies cannot be false. Thus, it remains to consider '$(A \in \mathbf{VER}) \wedge \neg(A \in \mathbf{TAUT})$'. It implies

(15) $A \Leftrightarrow \neg(A \in \mathbf{TAUT})$.

Furthermore, using (13), (**TS**), (**BI**) and the properties of material implication, we get

(16) $(\neg A \in \mathbf{VER}) \Leftrightarrow\ \vdash (A \Leftrightarrow \mathbf{1})$.

Since the right part of (16)—the formula $\vdash (A \Leftrightarrow \mathbf{1})$ is equivalent to $\vdash A$, we finally arrive at

(17) $(\neg A \in \mathbf{VER}) \Leftrightarrow\ \vdash A$.

This last assertion cannot be accepted—because if our logic is correct (sound), as it is—it does not prove falsehoods. Thus, we must reject (13). Observe that this conclusion depends essentially on (5), because (**TS**) prefixed by the provability sign allows us to pass from (6) to (13). The above argument demonstrates that if **T**-sentences are logical tautologies, the predicates 'is true' and 'is a tautology' have the same extension. However, this conclusion is counterintuitive since we are inclined to look at tautologies as a special case of truths, because not every true sentence belongs to logically true assertions. In other words, if A is a tautology, (13) holds for it and constitutes an instance of (**TS**). This is, however, a fairly trivial statement, because **T**-biconditionals as such do not distinguish between various kinds of truths. Thus, if the predicates 'is true' and 'is a tautology' are extensionally equivalent, the same applies to the predicates 'is false' and 'is a counter-tautology'. These facts motivate that (5), crucial to the entire presented argument, is untenable and should be rejected.

There is still another argument which shows that (12) entails that every sentences is logically determined (**LD**), i.e. is either a tautology or a counter-tautology (**CTAU**). Assume the following equivalences:

(18) $A \in \mathbf{LD} \Leftrightarrow A \in \mathbf{TAUT} \vee A \in \mathbf{CTAUT}$;
(19) $A \in \mathbf{TAUT} \Leftrightarrow \neg A \in \mathbf{CTAUT}$;
(20) $\neg(A \in \mathbf{VER}) \Leftrightarrow \neg A \in \mathbf{VER}$,

and in addition that A is true. These assumptions lead to

(21) (a) $A \in \mathbf{LD}$;

(b) $\neg A \in \mathbf{LD}$.

Assuming (**BI**) and $\neg A$ as true, we can prove (21). Since both of the assumptions pertaining to the logical value of A give the same result and exhaust all possibilities, we conclude that every sentence is logically true or logically false—that is logically determined.

Another argument, intended to show that considering **T**-equivalences as tautologies raises doubts, make use of semantics. If (5) is tautological, both its parts are satisfied by the same valuations. Consequently, the assertion

(22) For any valuation $\mathbf{v}$, $\mathbf{v}(A \in \mathbf{VER}) = \mathbf{v}(A)$),

is correct. On the other hand, since formula (5) as a tautology is satisfied by every valuation, it is equivalent to (22). However, this equivalence requires a further justification, because, due to the fact that the formulas A and $A \in \mathbf{VER}$ are syntactically different, we cannot assume in advance that the symbol **VER** is redundant in '$A \in \mathbf{VER}$'. The additional justification can only appeal to

(23) $(A \in \mathbf{VER} \Leftrightarrow A) \Leftrightarrow (A \Leftrightarrow A)$.

In fact, '$A \in \mathbf{VER} \Leftrightarrow A$' must be a tautology if (23) holds. Applying (9) and the equivalence $\vdash (A \Leftrightarrow \mathbf{1}) \Leftrightarrow A$, results in the right part of (21). Since all steps consist in operating on equivalences, they justify (21). Of course, we can use any other tautology instead of the formula $A \Leftrightarrow A$, but this change is of no importance, because we always use formulas equivalent to A on purely logical grounds, and not as a result of definitional replacement. The last restriction is crucial for considering (23) as equivalent to (22). Yet I do not think that Putnam would like to regard **T**-sentences as trivialities of the form $A \Leftrightarrow A$. Although Putnam is right in his claim that if **T**-equivalences are to be logical truths, their validity should be completely reducible to logical axioms and formal principles of the naming of expressions, his approach is the result of a decision about the status of **T**-sentences. Observe that Putnam's reasoning cannot be conducted under the assumptions related to **CT**, which require that **T**-equivalences are consequences of a truth-definition. In fact, we should rather speak about the metatheory of truth (I will denote it by the symbol **MT**). Although we can prove

(24) $\mathbf{MT} \vdash ((A \in \mathbf{VER} \Leftrightarrow A) \Leftrightarrow \mathbf{1})$,

this only implies that the formula is true in the models of **MT**, but not in all models. Thus, (23) provides no justification for regarding **T**-sentences as logical tautologies.

Contrary to the preceding arguments, the next (and the last I take into account) reason for rejecting the view that **T**-equivalences are tautological does not appeal to bivalence. The equivalence $\mathbf{T}A \Leftrightarrow A$ (this notation is more convenient in this context) has two implications, namely $\mathbf{T}A \Rightarrow A$, and $A \Rightarrow \mathbf{T}A$, as its components (see Chap. 4, Sect. 4.8). The former is frequently adopted as one of the axioms for the logic of truth, when truth functions as a modality, but the latter conditional is omitted or supplemented by additional constraints in order to eliminate **LP**. However, an additional reason can be given for rejecting the formula $A \Rightarrow \mathbf{T}A$ as universally valid. Suppose that we are working in the framework of a three-valued logic of Łukasiewicz. Take $v(A) = ½$. Clearly, '$v(A) = ½$' is true, but $\mathbf{T}A$ is false. This shows that the implication $A \Rightarrow \mathbf{T}A$ cannot be considered as a theorem of metalogic, although the formula $\mathbf{T}A \Rightarrow A$ still holds in our three-valued logic and its metatheory. If we accept the implication $\mathbf{T}A \Rightarrow A$ as tautological, but reject the logical validity of the conditional $A \Rightarrow \mathbf{T}A$, the formula $\mathbf{T}A \Leftrightarrow A$ cannot be considered as a logical theorem. I know of no other possible arguments that would allow us to interpret **T**-sentences as logical tautologies.

As I have noted already, Tarski claimed that **T**-equivalences should be provable at the level of **MT**. It can be done by the fixed point construction (see Chap. 8, Sect. 8.5). This observation resolves fully the issue of the status of **T**-equivalences from the mathematical point of view. On the other hand, philosophers have some reasons for regarding this point of view as merely partial, and seeking for further qualifications of **T**-sentences—for instance, in terms of tautologicity, analyticity or necessity. Although these concepts are subject to numerous fundamental controversies, their place in philosophical semantics, both formal and informal, appears well-established. Hence, it is difficult to imagine that discussions about the semantic status of **T**-equivalences, and of other formulas taken as important adequacy conditions in philosophical proposals could, dispense with the mentioned semantic notions. I will argue in the next paragraph that we might add something more explicit to the proposals by David and Künne (see above) concerning the necessary (conceptual) character of **T**-sentences.

The foregoing considerations suggest the following proposal by way of developing the ideas of David and Künne. We should avoid both of the characterized extremes, that is, the views of Putnam (**T**-equivalences) are tautologies of logic) and of Dummett (**T**-sentences are material equivalences). The proper solution must consist in a compromise which combines the thesis that **T**-biconditionals are stronger than material equivalences and the thesis that they are weaker than tautologies. The schematic outline of my proposal is as follows (see Woleński 2004g for details). Define semantically absolute analytic sentences as sentences which are true in all models. They coincide with logical truths of first-order logic. By the completeness theorem, semantically absolute analytic sentences are also absolute in the syntactic sense (provable from the empty class of assumptions). If **Th** is an axiomatic extralogical theory, its axioms and consequences form the set of sentences true in all its models. We can call them semantically relative analytic sentences of **Th**. This leads to the crucial conclusion that **T**-equivalences are semantically relative analytic sentences of **MT**. Note, however, that the set of truths of a given **Th** does not in general coincide with the semantically relative analytic sentences, because the theory can be incomplete; this is precisely the case of **MT**. Thus, we need to enlarge the concept of analyticity by adding pragmatically relative analytic sentences, that is, true in standard models—in $\mathbf{\Omega}$, in particular (see Chap. 8, Sects. 8.3 and 8.6). For instance, the equivalence "**AR** is consistent' is true if and only if **AR** is consistent' is true in the standard model of arithmetic. The connection between the property of being analytic and the property of being standard is justified by the fact that this property has its source in pragmatic decisions. The formula (4) in which 'snow is white' means that blood is red, can also be interpreted in terms of standard and non-standard models (in fact, I indirectly suggested that). This is because the sentence (1) holds in the standard model of our ordinary parlance, contrary to the equivalence (4), which becomes legitimate under an arguably non-standard interpretation.

According to the previous paragraph, David's "something more" or Künne's conceptual connections can be viewed as generated by appropriate analytic sentences. This observation suffices to reply to Dummett. Although **T**-sentences can be formalized by the sign $\Leftrightarrow$, they are strengthened material equivalences by being entailed by specific axioms. If we correlate various kinds of analyticity with corresponding sorts of necessity and apriority, in particular, absolute and relative (conditional), the extension of the proposed approach to **T**-equivalences becomes straightforward. In my opinion, it satisfies Halbach's claim that one should provide a theory of analyticity before demonstrating that **T**-equivalences are non-contingent. The view that **T**-biconditionals are relative analyticals concurs with Tarski's claim (see Convention **T**) that **STD** should logically entail such equivalences. This means that provability of **T**-biconditionals is a necessary condition of the correctness of the semantic definition of truth (this question is discussed in Pattterson 2006). I suspect (although I cannot illustrate this statement by a concrete textual illustration) that provability of **T**-sentences as a necessary condition of a correct truth-definition (Tarski addressed this condition to all truth-definitions proposed as satisfactory) is sometimes confused with the status of particular instances of **T**-scheme as necessities.

Some authors (see Black 1948, Haack 1978, pp. 100–101) argue that **CT** can be adopted also by other truth-theories, even very strange. Susan Haack considers the following example:

(25) A sentence is true if and only if it is asserted in the Bible,

and illustrates the issue by

(26) 'Warsaw was bombed in World War II' is true if and only if it is asserted in the Bible that Warsaw was bombed in World War II.

Haack says that if anyone remarks that the Bible does not say anything about bombings in World War II, the defender of (25) can reply that the relevant information is provided "in an obscure passage in Revelations", or (p. 101):

> [...] if he agrees that 'Warsaw was bombed in World War II', he will also, if he is wise, maintain the falsity of the right side of the scheme of the above instance of the schema. So, rather surprisingly, Tarski's material adequacy condition cannot be relied upon to be especially effective in ruling out bizarre truth definitions.

Clearly, if someone maintains that a given account of truth—for instance, that adduced by (25)—agrees with the intuition that truth consists in saying how things are, he could also argue that **CT** holds for the theory in question (for instance, see (**DG1III**) on James; he did not say anything about **T**-sentences). Haack's appeal to the wisdom is not convincing, because the person who accepts (26) but denies that the bombing of Warsaw in World War II is mentioned in Revelations, should also reject that Warsaw's bombing in 1939–1945 took place. In such a case, this person would use a model different from displaying how things are (were). The problem posed by (25) and (26) also consists in the fact that both are intensional, due to 'asserted that', whereas **T**-sentences are extensional. To conclude, I cannot agree with the assertion that **CT** fails "in ruling out bizarre truth-definition".

A more serious problem involves theories of truth associated with non-classical logic. As I argued in Chap. 4, Sect. 4.8, (**BI**) is stronger than (**TS**). Hence, the latter can be consistent with logic without (**BI**). For instance, nothing rules out to incorporating (**TS**) into many-valued logic, or logic with true value gaps, because we have that $\neg$(A is true) if and only $\neg A$. Although the left side of this equivalence is ambiguous, for the expression '$\neg$(A is true)' does not mean that A is false, the entire construction is defensible. On the other hand, (**SDT**) cannot be accepted since it entails (**BI**). Neil Tennant (see Tennant 1987, pp. 70–75) claims that **T**-sentences can accepted by the intuitionist as explicating meaning conditions in Davidson's sense (see Sect. 9.5), but it requires a modification of the sense of logical constants. Roughly speaking, sentences are true due to their acceptance conditions related to intuitionistic logic. This approach to truth does not generate (**TS**) in Tarski's sense, but other formulas (see Tennant 1997, p. 298)—similar to (**TS**) though not identical with it. A general form of such schemata is generated by

(27) A is true if and only if A satisfies the criteria of acceptance.

The exact meaning of (27) depends on how the criteria of acceptance are defined. Nothing more can be said—particularly on the relation between concrete instantiations of (27) and (**TS**)—without knowing details of how acceptance is to be understood, even if (27) is supplemented by the constraint that the logic of acceptance must be effective, constructive, recursive, etc.

9.4 Truth and Meaning

Since **STT** concerns interpreted languages, it assumes that if **L** is such a language, its expressions have meaning (see Chap. 7, Sect. 7.5 and **DG8**). Davidson (see Davidson 1967) reversed the direction and tried to explain the concept of meaning via the concept of truth, using **T**-equivalences as the basic tool. Thus, If **L** is a language (natural or formalized), its theory of meaning is captured by **T**-sentences. To put it another way, we axiomatize the theory of meaning for **L** by instantiating (**TS**) for any sentence $A \in$ **L**. According to Davidson (see Davidson 1970, p. 60), the phrase 'is true if and only if can be read as 'means that,' and (**TS**) becomes

(28) 'A' means that A^*.

For example, the sentence

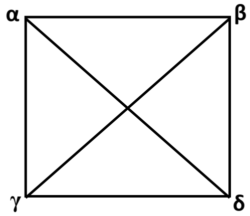
(29) 'London is a city' is true if and only if London is a city,

which is an instance of (**TS**), can be interpreted as

(30) 'London is a city' means that London is a city.

Roughly speaking, the meaning of a sentence is explained by its truth-conditions.

I will argue that this account is untenable, because the properties of ‘it is true’ are different than the properties of ‘means that’. Truth and meaning can be considered as modalities. Consider truth first (see Chap. 4, Sect. 4.8 for a more detailed account). We transform ‘A is true’ into ‘it is true that A’ (I neglect here the **L/ML** distinction and assume that dangerous sentences, like the Liar, are excluded). Thus, ‘it is true that’ functions as a monadic modal operator acting on sentences. The basic logical facts about ‘it is true that A’ (symbolically, $\mathbf{T}(A)$) are displayed by diagram (**D2**) (it is a part of the diagram (**D1**) from Chap. 4, Sect. 4.8):



where we interpret α as $\mathbf{T}(A)$, β as $\mathbf{T}(\neg A)$, γ as $\neg\mathbf{T}(\neg A)$, δ as $\mathbf{T}(\neg A)$. Recall the following principles from Chap. 4, Sect. 4.8:

(31)

(a) $\neg(\alpha \wedge \beta)$ ($\mathbf{T}(A)$, $\mathbf{T}(\neg A)$ are contraries);
(b) $\alpha \Rightarrow \gamma$ ($\mathbf{T}(A)$ entails $\neg\mathbf{T}(\neg A)$);
(c) $\beta \Rightarrow \delta$ ($\mathbf{T}(\neg A)$ entails $\neg\mathbf{T}(A)$);
(d) $\alpha \Leftrightarrow \neg\delta$ ($\mathbf{T}(A)$, $\neg\mathbf{T}(A)$ are contradictories);
(e) $\beta \Leftrightarrow \neg\gamma$ ($\mathbf{T}(\neg A)$, $\neg\mathbf{T}(\neg A)$ are contradictories);
(f) $\gamma \vee \delta$ ($\neg\mathbf{T}(\neg A)$, $\neg\mathbf{T}(A)$ are complementaries).

Since the formula $\mathbf{Mean}(A)$ (it means that) satisfies the principles listed in (3), so far we have full symmetry between truth and meaning. However, matters become more complicated if we ask for further rules. In particular, the problem arises how both operators behave with respect to negation and non-modalized sentences. If we consider these additional aspects, the question of the relation of **Mean** and **F** (falsehood) immediately shows up. Diagram (**D2**) suggests nothing for these issues.

Since Davidson’s approach uses **T**-sentences, we have:

(32) $(\mathbf{T}(A) \Leftrightarrow A) \Leftrightarrow (\neg\mathbf{T}(A) \Leftrightarrow \neg A) \Leftrightarrow (\mathbf{T}(\neg A) \Leftrightarrow \neg A) \Leftrightarrow (\mathbf{F}(A) \Leftrightarrow \neg A)$.

This assertion means that if **T** commutes fully with negation and ‘not-truth’ (it is not the case that … is true), then ‘truth not’ and ‘falsehood’ are exactly co-extensional; in particular, not-truth is falsehood. Things are different with **Mean**, because we cannot equate $\neg\mathbf{Mean}(A)$ with $\mathbf{Mean}(\neg A)$, that is, ‘it is not the case that ‘A’ means A’ and “A’ means not-A’. Clearly, the sentence ‘it is not the case that ‘London is a city’ means that London is a river’ is not equivalent to “London is a city’ means that it is not the case, that London is a river’. In fact, the former entails the latter, but the reverse implication does not hold. **T**-sentences establish a very strong connection between **T**(A) and A, namely that both are equivalent, but we cannot accept neither (a) **Mean**(A) $\Rightarrow A$, nor (b) $A \Rightarrow$ **Mean**(A). The reason for rejecting (a) is obvious, for the sentence '2 + 2 = 5' means that 2 + 2 = 5—but is false. Hence, we cannot assert of the **T**-sentence that it is related to (a). At the first glance, (b) seems plausible, because one might argue that, if **T**('A') $\Leftrightarrow$ (A) holds, then it is required that 'A' means A. Nevertheless, there are cases that are at odds with this claim. If A is an arithmetical sentence, true in standard as well as non-standard models, its meaning depends on some additional constraints, which are not captured by (b) or related T-equivalences. The differences between **T**(A) and **Mean**(A) have a very simple explanation. Truth, at least under **STT**, is purely extensional, but meaning cannot be considered this way.

(**DG5**) I limit my considerations in this section to the logical aspects of Davidson's project. In fact, he tried to establish some empirical conditions for asserting that 'A means that'. I do not enter into the problem of whether the theory of meaning for **L** as an empirical theory based on **T**-sentences liquidates the fundamental gap between intensionality of **Mean** and the extensionality of **T**. I presented my arguments in a talk on the views of Davidson at the conference in Kazimierz Dolny (Poland), in 1995 (see also Woleński 2007a). Davidson was present and replied by pointing out that his intention was to explain an important conceptual link between two notions, namely truth and meaning. However, my argument did not question that such a link exists and is important. My principal aim was to show that since the nature of truth is fairly different—modulo the extensionality/intensionality feature—than the nature of meaning, More specifically, the concept of truth as defined by Tarski is extensional, but the concept of meaning—intensional. Davidson's attempt fails at least insofar as logical problems are concerned, because the reduction of intensionality to extensionality is problematic, if possible at all.►

(**DG6**) The questions pertaining to Tarski's attitude toward Davidson's idea that the meaning of a sentence can be explained by its truth-conditions is a historical one. When I met Davidson in Berkeley in 1989, I asked him whether he discussed his views with Tarski. He answered that he did not. According to Hintikka (personal communication), Tarski was very critical about Davidson's approach to meaning.►

Some authors (Künne 2003, pp. 333–350; A) introduce the meaning parameter into (**TS**) directly. This move leads to ('means **m**' refers to the meaning of a sentence)

(33) 'A' is true if and only if ('A' means **m**) $\wedge$ A.

By contraposition, we obtain

(34) 'A' is false if and only if ('A' does not mean **m**) $\vee$ $\neg A$.

A concrete illustration of (6) is

(35) 'Snow is white' is false if and only if 'snow is white' does not mean that snow is white or snow is not white.

Assume that ‘snow is white’ means that grass is green (**m**(‘snow is white’) = snow is white. The right part of (35) is true owing to the truth of “‘Snow is not white’ does not mean that snow is white’. Consequently, the sentence ‘Snow is white’ is false, independently of whether grass is green or not. Thus, (34) does not offer a satisfactory account of ‘is false’. A more sophisticated account was proposed in Prior 1971, p. 104 (see also Künne 2003, p. 347). Transform (5) into

(36) ‘A’ is true $\Leftrightarrow \exists x(x = \mathbf{Mean}(A) \wedge A$.

However, this modification does not help because it results in

(37) ‘A’ is false $\Leftrightarrow \forall x(x \neq \mathbf{Mean}(A) \vee \neg A$,

because it admits that a sentence if false, if it is meaningless. If (see Hugly, Sayward 1996, p. 356), we adopt

(38) ‘A’ false $\Leftrightarrow \forall x(x = \mathbf{Mean}(A) \wedge \neg A$.

Due to this definition of falsity, the sentences ‘A is false’ and ‘A is true’ are not mutual negations (see also Sect. 9.7).

Inaccuracies related to (34), (37) and (38) can be overcome by presupposing that the expressions of **L** have meaning. In other words, the statement stating that A means that **m**, is not a part of the definition of truth. In fact, **STT** takes this path, made explicitly by Tarski (see Chap. 7, Sect. 7.4)—that the expressions of **L** are equipped with meaning. Although stating that meanings occurring in **L** are explained in **ML** is not trivial, it does not contribute very much to the question ‘What is meaning?’—absolutely basic for the philosophy of language. On the other hand, one can say that most theories of meaning can be accommodated by **STT**, similarly as in the case of truth-theories associated with (33) and (36). In fact, **STT** does not depend essentially on any particular theory of meaning. However, its important factor consists in accepting that if E is an expression of **L**, it has the standard (ordinary—in Ryle’s sense; see (**DG5VI**)) meaning exhibited by the resources of **ML**. If E has (or may have) a non-standard meaning, this fact also sends us to the metalanguage of **L**. In other words, the assertion that ‘A’ is true basically depends on whether we take the meaning of A as standard or not. This conclusion allows us to meet Putnam’s objection (see Putnam 1983; see Woleński 2001 for criticism of Putnam’s view) that **STT** is fundamentally asemantic, because it makes no distinction between a sentence A as signifying, for instance, that snow is white, or the same sentence as signifying that grass is green. On the contrary, **STT** makes a direct reference to an interpretation $\mathfrak{I}$.

Another argument that **STT** assumes a wrong semantics can be found in Etchemendy 1990, p. 15. This author claims that a good semantic theory should help in answering whether the sentence ‘snow is white’ would be true in the case of snow being black. John Etchemendy proposes to consider truth-conditions for the sentences ‘snow is white’ and ‘roses are red’. According to **PC**, we have just four well-known cases, but there is the case in which both sentences are true. However, this treatment of truth-tables is not correct, because valuations in **PC** are mappings from the set of variables to the set of truth-values and do not pertain to sentences having a fixed meaning under a specific interpretation $\mathfrak{I}$. Etchemendy claims that a proper semantics should be representational in the sense that it displays the representation of the world. This suggestion is perhaps very rational and promising, but **STT** has nothing to do with representations of reality. In particular, **T**-sentences do not say that concrete sentences are true or false, but establish conditions of being true in models. That's all.

(**DG7**) In **DG14VII**, I pointed out Tarski's preference for a definition of truth over the axiomatic method of explaining the predicate 'is true'. The related problem of physicalism was discussed in Field 1972 and Kirkham 1993. Richard Kirkham argues that Tarski accepted physicalism as the most proper philosophical foundation of semantics. However, as I mentioned in Chap. 6, Sect. 6.4, Tarski's relation to philosophy was deliberately ambiguous and his declaration, in fact quite marginal in Tarski 1936, must be taken *cum grano salis*, and cannot be regarded as an interpretative guide. According to Hartry Field, **STT** should be "physicalized" by the concept of primitive denotation. This suggestion requires a radical change of metamathematics and cannot be discussed here.►

(**DG8**) Let me return once again to Tarski's claim (see Chap. 7) that we ascribe concrete and intelligible meanings to the signs which occur in the considered (interpreted languages. Although Tarski consequently avoided (see **DG10VII**) explaining what meaning is, the problem of sources of intelligibility of linguistic expressions remains. A partial answer can be propose via the concept of standard model (see Chap. 8, Sect. 8.6). It is not accident that, as I noted in Chap. 8, Sect. 8.6 such models are called 'intended', that is, capturing our intuitions, for instance on natural numbers. Thus, we can say that standard models are inherently associated with intelligible meaning; on this occasion I remind once again Ryle's ideas that the non-ordinary terms also have the standard use (see **DG7VI**). On the other hand, due to metamathamatical considerations about ω-concepts, we should not say that non-standard models are not intelligible. In particular, qualifying that a model **M** is non-standard assumes that we are able to pick up the standard structure. Although sometimes it is highly controversial (see the case of set theory mentioned in Chap. 8, Sect. 8.6), the scientific and commonsense practice suggests that ordinary language provides resources for accounting what is standard or non-standard, at least for the first sight. Such qualifications are always revisable, but their basic role cannot be denied. Due to this observation, ordinary language (in Ryle's sense) is the fundamental source, although not ultimate, of intelligibility of meaning, even in the case of formalized languages. These observations agree with Leśniewski's maxim (see Introduction) that logic is a formal exposition of intuition as well as with Tarski's central maxim (see Chap. 6, Sect. 6.6, **DG8VI**, **DG19VII**) that the problem of truth is meaningless for purely formal languages. Generally speaking, informal metalanguages are prior to formalized languages. The intelligibility of meaning in nothing mysterious in this perspective.►

9.5 Is SDT Absolute or Relative?

Various authors qualify differently **SDT** form the point of view of the absolutism/relativism controversy. For instance, Popper maintains (see Popper 1972, p. 46) that Tarski's theory is absolute, because objective. Davidson (see Davidson 1973, p. 68) argues that truth for interpreted languages is absolute, but relative for non-interpreted languages. Carnap (see Carnap 1942, p. 240, Carnap 1947, pp. 93–94; a similar view is in Pap 1954) says that if truth is attributed to propositions, it is absolute, but if to sentences—it is relative; this view entails that **SDT** is relative if sentences are taken as truth-bearers. For Przełęcki (see Przełęcki 1974), **SDT** is relative if defined for a possible interpretation, but absolute for the standard model (the real world). According to Susan Haack (see Haack 1978, pp. 114–115), Tarski's theory is relative due to the fact that it defines truth relative to a language **L**, but languages can have different interpretations. Tarski himself addressed the problem of relativism only once, namely in 1933, p. 199. He suggested that truth in a domain that is a subset of all objects is relative. Consequently, *a contrario*, truth in the entire the universe is absolute. However, this suggestion was marginal and ignored in Tarski's further analysis. The above survey shows that many intuitions are attached to the concepts of absolute and relative truth. Some views can be immediately regarded from the point of view adopted in this book. Since I consider sentences as truth-bearers, the distinction of propositions and sentences is irrelevant (I do not say that it is not relevant at all; see Simons 2003 for arguments that considering sentences as truth-bearers does not force relativism). Similarly, since **SDT** assumes that **L** is interpreted, we do not need to take into account truth for non-interpreted languages.

I will take Kokoszyńska's analysis as the starting point (other approaches are proposed for instance, in Dawson-Galle 1998, Kölbel 2002, O'Grady 2002, García-Carpintero, Kölbel 2008). She (see Kokoszyńska 1936) interpreted the classical truth-definition as absolute and instantiated by **SDF**. In Kokoszyńska 1948 and Kokoszyńska 1951, we find a penetrating discussion of the relativity of truth (unfortunately her papers are unknown or ignored in contemporary books on the aletheiological relativism) inspired by Twardowski (see Chap. 3, Sect. 3.7.1). According to Kokoszyńska, the predicate 'is true' is incomplete as such, and can be completed in various ways—for instance by reference to some definite circumstances. If a sentence A has a fixed meaning, it is relatively true if there exist circumstances $\mathbf{C}$ and $\mathbf{C}'$ such that A is true in $\mathbf{C}$, but $\neg A$—in $\mathbf{C}'$ (or conversely). Such relativism is proper. It can be either radical (if every sentence is relatively true) or moderate (if relativity applies to some sentences only). On the other hand, improper relativism admits that sentences are true in some models, but false in others. Kokoszyńska's analysis implies that proper relativism agrees with the temporal change of logical values in the same model. As far as many-valued logic is concerned, it falls under relativism unless we assume that only a change of truths into non-truths entails relativity. Although Kokoszyńska tried to make Twardowski's intuitions precise, but she also followed the Kotarbiński/Leśniewski debate (see Chap. 3, Sect. 3.7.2) and the discussions around many-valued logic (see Chap. 3, Sect. 3.7.3), although her main task consisted in preparing the absolutist interpretation of **SDT**. In other words, Kokoszyńska intended to combine Tarski's ideas with that of Twardowski.

My further analysis concerns the understanding of the absoluteness of truth as related to its eternality and sempiternality (see Chap. 3, Sect. 3.7.2; I recall some definitions from this fragment but I omit bibliographical references provided earlier). The equivalence

(39) Truth is absolute if and only if truth is eternal,

expresses the weak truth-absoluteness thesis (**WTAT**). The strong truth-absoluteness thesis (**STAT**) says that truth is absolute if and only if truth is eternal and sempiternal. I states that

(40) A sentence is true at a moment t if and only if it is true at any other moment t'.

This statement has two parts:

(41) (a) (eternality, **ET**) If a sentence is true at t, it is also true at any $t' \geq t$;

(b) (sempiternality, **ST**) If a sentence is true at t, it is also true at any $t' \leq t$;

Consider a concrete example, namely the sentence

(#) I will go to Warsaw on April 26, 2018, written on September 3, 2017.

Assume also that I will stay in Warsaw on April 26, 2018. Thus, (#) is true on t = April 26, 2018. Consequently, we have:

(42) (a) ($\mathbf{ET^{(\#)}}$) If (#) is true at t, it is also true at any $t_1 \geq t$;

(b) ($\mathbf{ST^{(\#)}}$) If () is true at t = April 26, 2018, it is also true at any $t_1 \leq t$.

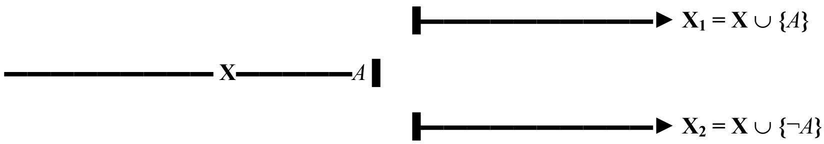
Clearly, if we take three-valued logic as the basis, A is valued by the neutrum and becomes an indefinite sentence. Consequently, although $v(A) = ½$, (42a) holds (it will be true or false at t_1), (42b) cannot be evaluated, because its sempiternality is ruled out. Combining (40)–(42) with distinctions introduced by Kokoszyńska, eternality without sempiternality characterizes moderate relativism as far as temporal coordinates of sentences are concerned, provided (see below) that truth is stable in the sense that what was true, remains such for ever.

Leśniewski offered a proof that, assuming the law of non-contradiction, **ET** and **ST** are equivalent, that is, **STAT** holds. His argument is as follows (this reconstruction is an extension of remarks in Chap. 3, Sect. 3.2.2). Assume that a sentence (a) 'S is P' is true, but not sempiternally true. Thus, there is a time t in which (b) is not true, although it is true at some time $t_1 \geq t$. This assertion implies that (c) 'S is not P' is true at t. Leśniewski uses the principle of non-contradiction at this point and concludes that since (c) is always false, it is also not true at t. Thus, (b) must be true at t, contrary to our initial assumption. Hence—an inconsistency. Leśniewski's proof that eternality entails sempiternality proceeds analogously. Leśniewski's reasoning was informal. One of its crucial steps can be questioned, namely the one in which Leśniewski assumes that falsehood is sempiternal. One might observe that if we assume that the property of 'being false' obtains sempiternally, the same applies automatically to truth. Fortunately, the outlined proof has its contemporary setting, which does not require any appeal to the sempiternality of falsehood. Assume that (i) **X** is a true set of sentences about the past; (ii) *A* is a future contingent; (iii) *A* is independent of **X**; assumption (iii) is necessary, because the dependence of *A* would imply **STAT**. Assumption (i) implies that **X** is consistent and thereby possesses a model **M**. Due to (iii), the sets $\mathbf{X_1} = \mathbf{X_1} \cup \{A\}$ and $\mathbf{X_1} = \mathbf{X_2} \cup \{\neg A\}$ are consistent and have models, say, $\mathbf{M_1}$ and $\mathbf{M_2}$. We can think about $\mathbf{M_1}$ and $\mathbf{M_2}$ as models of maximally consistent sets of sentences; this is perhaps the best intuition related to the concept of possible worlds. In virtue of Lindenbaum's theorem on maximalization, every consistent set of sentences has its maximal extension. Thus, the set **X** has at least two such extensions those that contain the sets $\mathbf{X_1}$ and $\mathbf{X_2}$, respectively. Similarly, the two models $\mathbf{M_1}$ and $\mathbf{M_2}$ extend the model **M**. More technically, **M** is a submodel of $\mathbf{M_1}$ and $\mathbf{M_2}$. Clearly, sentences true in **M** remain true in $\mathbf{M_1}$ and $\mathbf{M_2}$.

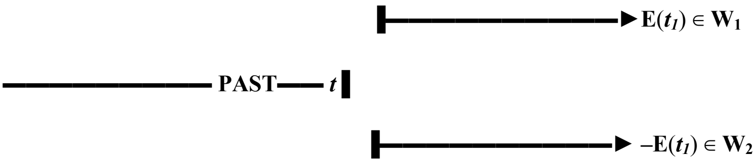
Another construction which justifies (c) is as follows. We define (see Asser 1972, pp. 168–169) the concept of branchability by (**X**—a set of formulas):

(43) (a) **X** branches at a formula *A* if and only if the sets $\mathbf{X_1} \cup \{A\}$ and $\mathbf{X_2} \cup \{\neg A\}$ are consistent;
(b) **X** is branchable if and only there is a formula *A* at which **X** branches;
(c) **X** is branchable if and only if **X** is a consistent and incomplete set of sentences.

Since we have no reason to assume that **X** is a complete set of sentences, we say that **X** branches at *A*. This is displayed by the diagram (**Δ**):



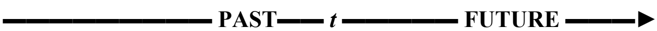
An ontological interpretation is also possible. Let *A* be a sentence uttered at a moment fixed as the present (denoted by *t*) and refers to a future event **E**, which will happen at t_1 ($\mathbf{E}(t_1)$) or will not happen at t_1 ($-\mathbf{E}(t_1)$); the segment indicated by **PAST** covers everything that happened before *t* (including this moment). We transform the diagram (**Δ**) into (**Δ′**)



$\mathbf{W_1}$ and $\mathbf{W_2}$ are possible worlds, that is, possible continuations of **PAST** as their initial segment. To put it differently, the worlds $\mathbf{W_1}$ and $\mathbf{W_2}$ augment **PAST**, but in radically different (or ontologically inconsistent) ways, because the former validates *A*, but the latter verifies $\neg A$.

The status of **M** as a submodel of $\mathbf{M_1}$ and $\mathbf{M_2}$ justifies **ET**. In fact, if *A* will be true at the instant t_1 later than *t*, its truth becomes stable for ever. Thus, **ET** is automatically valid. For the sake of further argument, we can consider the value of *A* at any temporal point *t* earlier that at t_1, not necessarily at the moment in which this sentence is (was) really uttered. Assume that *A* is eternally true, that is, if *A* is true at *t*, it remains true at any $t_1 \geq t$. This means that *A* is true in $\mathbf{M_1}$. However, the sentence in question cannot be false in any submodel of $\mathbf{M_1}$, in particular in **M**. Hence, we immediately conclude that if *A* is eternally true, it is sempiternally true as well. The same line of reasoning applied to the assumption that $\neg A$ is true in $\mathbf{M_2}$ implies that if *A* is false in $\mathbf{M_2}$, it is also false in **M**. Thus, the both logical values are perfectly stable over time, **ET** and **ST** are equivalent, and **STAT** finds its support. We have now, exact tools for comparing many-valuedness, (**BI**), and truth-absoluteness. The argument for **STAT** uses essentially (**BI**). The converse dependence holds as well. Hence, (**BI**) and **STAT** are equivalent, at least in the model-theoretic semantics. Nonetheless, if truth-absoluteness is reduced to eternality, **(BI)** has no explicit connection with **ET**. One can assume that any indefinite sentence becomes true or false sooner or later, but the conjecture that at least some indefinite sentences remain such for eternity is also possible. In both cases, however, **ET** can be defended. Thus, Łukasiewicz (see Chap. 3, Sect. 3.7.3) could consider truth as absolute, but it is possible provided that this property would be understood in a weaker sense, that is, according to **WTAT**. On the other hand, **SDT** is strongly absolute, because it satisfies **STAT**.

Many philosophers, Łukasiewicz among them, accept the eternality of truth as uncontroversial, but complain about sempiternality. The argument for the logical equivalence of (**ET**) and (**ST**) shows that something is wrong in this view, because if sempiternality is felt as a fairly non-intuitive property, the same should apply to eternality. I only indicate this problem briefly. Its full analysis requires taking into account ontological questions related to determinism and its various forms. For example, radical determinism can be displayed by the diagram ($\mathbf{\Delta}''$)



This line symbolizes the uniform "flow" of reality from the past though the present moment t to the future. This flow can be organized by the strict (one-to-one) causal connection. In fact, Łukasiewicz argued that sempiternality is implied by **MEM** plus causality. However, the outlined argument shows that **STAT** can be defended on purely logical grounds and without an appeal to ontology. Thus, objections against sempiternality as an intuitive property of truth seem to emerge as a result of a conflation of logic with ontology that is not a rarity in the history of philosophy.

The mentioned conflation occurs in logical determinism (**LD**)—the view that classical logic entails radical determinism (**RD**). The term 'logical determinism' historically (more accurately, its German counterpart, that is, *logische Determinismus*) was introduced by Moritz Schlick. He spoke about the paradox of logical determinism paradox a emerging in a conflict between logic and the view, shared by many philosophers, that the future is open, not ready now, etc. Schlick's formulation is this (Schlick 1931, p. 202).

> […] the principles of contradiction and excluded middle would not rank as statements about future states-of-affairs if determinism did not prevail. In fact, so Aristotle argued, if determinism is true, and if the future, therefore, is not already laid down and *determined* now, than it seems as if the proposition "Event *E* will take place tomorrow' could today be neither true not false.

This problem was also discussed by Waismann who, besides 'logical determinism', also used the expressions 'logical predestination' or 'logically preordained' (Waismann 1956; Waismann 1965, pp. 27–34; Waismann 1976 (written in 1939, pp. 59–68); the following quotation is taken from (Waismann 1956, pp. 8–9)):

> I shall single out for discussion [of the nature of philosophical problems – J. W.] the question […] whether the law of excluded middle, when it refers to the statements in the future tense, forces us into a sort of logical Predestination. A typical argument is this. If it is true now that I shall do a certain thing tomorrow, say jump into the Thames, then no matter how fiercely I resist, strike out with hands and feet like a madman, when the day comes I cannot help jumping into the water; whereas, if this prediction is false now, then whatever efforts I may make, however many times I may nerve and brace myself, look down at the water and say to myself, 'One, two, three' – it is impossible for me to spring. Yet that the prediction is either true or false is itself a necessary truth, asserted by the law of excluded middle. From this the startling consequence seems to follow that is already now decided what I shall to do tomorrow, that indeed the entire future is somehow fixed, logically preordained.

There is an interesting difference between Schlick and Waismann (they did not accept **LD**) as to which the effect which logical principles are responsible for logical determinism. Schlick speaks about the laws of contradiction and excluded middle, but Waismann only of the second.

The principle (**BI**) is (see Chap. 4, Sect. 4.6) the conjunction of the excluded middle and the rule of non-contradiction, both understood metalogically. However, if we tacitly assume that we work in a consistent system, the former ((**EM**)) suffices for arguments. Let the symbol $\blacklozenge A$ means that A is contingent (= not necessary and not impossible). Thus, the contingency of A is defined by

(44) $\blacklozenge A =^{df} \Diamond(A) \wedge \Diamond(\neg A)$.

Accordingly, sentence (#) (I will go to Warsaw on April 26, 2018) has the form $\blacklozenge A$ and is read as 'it is contingent that I will be in Warsaw on April 26, 2108'. Using (40) with respect to $\Diamond A$, we obtain

(45) A sentence of the form $\Diamond A(t')$' is true at t if and only it is true at any $t'' \neq t$,

which asserts that the sentence $\Diamond(A(t')$ is sempiternally true. Having the basic intuitions specified, we can give the following explanations of logical determinism:

(46) (a) It is logically necessary that worlds with the same past and the same natural laws have the same future (a characterization of **LD**);

(b) **(EM)** implies **LD**;

(c) **STAT** implies **LD**;

(**DG8**) (7a) is equivalent to the following statements (i) the class of worlds with the same past is unique, that is, there are no different possible worlds; (ii) the class of possible future worlds = **{**the future of the actual world**}**—this means that the actual world continues uniquely from the past to the future; (iii) logic excludes that 'it is possible that A and it is possible that not-*A*' is a factor of 'it will be *A*', where *A* is neither tautological nor contradictory. I include (46c) since it characterizes **STT** and it is interesting how our understanding of truth is related to well-known ontological issues.►

I shall argue that all the mentioned versions of **LD** are not derivable from logic itself. However, this does not mean that I opt for logical indeterminism, that is, a negation of **LD** in one of its formulations. And I shall argue that the issue of determinism/indeterminism is independent of pure logic. Thus, the truth of future contingencies as defined in (44) is combined with adding (42) is consistent with determinism as well as indeterminism. Clearly, it is a very instructive example that logic does not distinguish extralogical matters (see Chap. 5(31)).

I shall once again use diagram **(D1)** from Chap. 4, Sect. 4.8 in the interpretation (since we discuss **LD**, all following formulas can be read logically or ontologically): α as A is determined, $\mathbf{D}(A)$; β as $\neg A$ is determined, $\mathbf{D}(\neg A)$; γ as $\neg A$ is not determined, $\neg\mathbf{D}(\neg A)$; δ as $\neg(A$ is determined), $\neg\mathbf{D}(A)$; ε as A is determined or $\neg A$ is determined ($\alpha \vee \beta$; $\mathbf{D}(A) \vee \mathbf{D}(\neg A)$; ζ as ($\neg A$ is determined) $\wedge$ $\neg(A$ is determined) ($\gamma \wedge \delta$; $\neg\mathbf{D}(\neg A) \wedge \neg\mathbf{D}(A)$; A is contingent); κ as A (it is true that A); λ as $\neg A$ (it is true that $\neg A$). From the rules of (**D1**) we obtain (among others):

(47) (a) $\mathbf{D}(A) \Rightarrow \Diamond A$;
(b) $\mathbf{D}(A) \Rightarrow \mathbf{T}(A)$;
(c) $\mathbf{T}(A) \Rightarrow \Diamond(A)$;
(d) $\blacklozenge(A) \Rightarrow \Diamond(A)$.

The converses of (47a)–(47a) do not hold on purely logical grounds, if $\mathbf{\Delta}$ displays our logic. Furthermore, we define (recall that A is neither tautological nor contradictory):

(48) (a) radical determinism (**RD**): $\forall A(\varepsilon)$, $\forall A(\alpha \vee \beta)$, $\forall A(\mathrm{D}(A) \vee \mathbf{D}(\neg A))$;

(b) radical indeterminism (**RI**): $\forall A(\zeta)$, $\forall A(\gamma \wedge \delta)$, $\forall A(\neg \mathrm{D}(A) \vee \neg \mathbf{D}(\neg A))$;

(c) moderate determinism (**MD**): $\exists A(\alpha \vee \beta) \wedge \exists A(\gamma \wedge \delta)$;

$$\exists A(\mathrm{D}(A) \vee \mathbf{D}(\neg A)) \wedge \exists A(\neg \mathbf{D}(A) \vee \neg \mathbf{D}(\neg A));$$

(d) moderate indeterminism (**MI**): $\exists A(\zeta) \wedge \exists A(\varepsilon)$, $\exists A((\gamma \wedge \delta) \wedge \exists A(\alpha \vee \beta)$;

$$\exists A(\neg \mathbf{D}(A) \wedge \neg \mathbf{D}(\neg A)) \wedge \exists A(\mathbf{D}(A) \vee \mathbf{D}(\neg A));$$

(e) minimal determinism (**DM**): $\exists A(\alpha)$, $\exists A\ \mathbf{D}(A)$;

(f) minimal indeterminism (**IM**): $\exists A(\zeta)$, $\exists A(\neg \mathbf{D}(A \wedge \neg \mathbf{D}(\neg A))$;

Intuitively, **RD** says that everything (in the world—I will omit this clause in further explanations) is determined, **RI**—that everything is not determined, **MD**—that something is determined and something else is not, **MI**—that something is not determined and something else is determined, **DM**—that at least something is determined, and **IM**—that at least something is not determined; **MD** and **MI** are logically not distinguishable, but can differ in the distribution of what is determined and what is not (hence, I introduced a different succession in the respective formulas, but this has no logical relevance). $\mathbf{D^f}$ entails **DM** and $\mathbf{I^f}$ entails **IM**, **DM** follows from **RD** and **MD**, **IM** follows from **RI** and **MI**. No statement of (48a)–(48f) expresses a logical truth and thereby cannot be entailed by (**EM**). Hence, we conclude that the existence of possible events is consistent with **RD**, **RI**, **MD**, **MI**, **DM** and **IM**, that is, with all kinds of determinism and indeterminism listed in (48). Also **LD** entails truth (see 47b), but if *A* is true, it does not mean that the event (fact, state of affairs, etc.) described by this sentence is determined. Similarly, our logic (see (47c) neither implies that there are necessary truth nor that some truths are contingent. To conclude, neither first-order logic nor the logic of truth as displayed by diagram (**D1**) justifies logical determinism as an ontological theory.

I shall illustrate the last (more precisely, its first part) conclusion by the following example. Observe that **RD** is traditionally expressed by the sentence 'everything is necessary'. This statement can be formalized as (I will use the box $\Box$ instead the letter **D**; I remind that necessity and other modalities can be interpreted ontologically that is applied to facts, events, etc.).

(49) $\forall A(\Box A \vee \Box \neg A)$.

In order to prove **LD** exlusively by the devices of logic (in this case modal logic) as the only means of inference from a theorem of logic as the only premise, the box should be interpreted as logical necessity (this means that logical and ontological modalities are equivalent; I do not enter into a discussion of this identification). The form of (49) suggests an appeal to (**EM)** as the starting point of an argument. Thus, assume this principle, that is, the formula $A \vee \neg A$. Apply necessitation to derive (@) $\Box(A \vee \neg A)$. However, (@) does not imply (49) (the converse implication holds). This fact, very often invoked (for instance, see Wessel 1999, pp. 151–152), does not depend on the logical or metalogical understanding of (**EM**). Thus, deriving **LD** from (10) is a simple logical mistake, because the latter is not a logical rule. In particular, it does not express (**EM**). In my opinion, this observation closes the issue of how classical logic is related to determinism. Once again, the answer is: in no way! To put it another way, any direct derivation of determinism (the same applies indeterminism) from pure logic consists in the mentioned conflating the latter with ontology.

(**DG9**) I do not claim that the list in (48) is exhaustive (see Earman 1986 for an elementary introduction to determinism and its various problems). For example, I entirely omitted the problem of (theological) fatalism and various questions related to the concept of causality as well as so-called statistical determinism or deterministic chaos. Also the relation of logic and determinism is more complicated than it was presentend in my remarks (see Vuillemin 1996, and Williamson 2013 for a more extensive analysis). Omissions and simplifications in my presentation of ontological questions can be explained by pointing out that the concept of truth as defined in **STT** is the main target of my considerations.►

What about **STAT** and **RD**? The answer can be obtained by branchability. Consider ($\boldsymbol{\Delta}'$) once again. We have two models $\mathbf{M_1}$ and $\mathbf{M_2}$ as extensions of **M**. Intuitively, **M** = **PAST**, but $\mathbf{M_1}$, $\mathbf{M_2}$—its future extensions; in other words, **PAST** is the initial segment of $\mathbf{M_1}$ and $\mathbf{M_2}$. However, any sentence true in **M** is true in its future extensions. As far as the issue of $\mathbf{M_1}$-truths and $\mathbf{M_2}$-truths is concerned, *A* is true in (this is only a picturesque notation) $\mathbf{M_1} - \mathbf{M}$; it must be true in **M** as well. Consequently, $\neg A$ is false in $\mathbf{M_2}$, and a fortiori—in **M**. Similar argument applies to $\mathbf{M_2}$-truths. In general, there is no way to pass from the situation described in ($\boldsymbol{\Delta}'$) to the world represented by ($\boldsymbol{\Delta}''$) (or conversely by employing pure logic). If we identify ($\boldsymbol{\Delta}''$) with a description of radically deterministic reality, logic does not guarantee generate that this picture is correct or not. In order to justify or reject ontological models of reality, we need something more that logic—for instance, the principle of causality or its more or less radical denial. **STAT** a characteristic feature of **STT**, as applied to sentences and their logical values, justifies neither determinism nor any other view mentioned in (48). Thus, sempiternality of truth has no fatalistic consequences. I consider this fact as a very interesting aspect of **STT** from the philosophical point of view.

(**DG10**) Metamathematical results suggest that Kokoszyńska's distinction should be modified in some way. Take **AR** as an example. Its axioms (and their logical consequences) are true in all models. On the other hand, the undecidable sentences are true in some models, but false in others; they are branching points in the sense of (43). Using, Kokoszyńska's language they are relatively true in the improper sense. Yet if *A* is true in a model **M**, it cannot be false in this model. This property characterizes all sentences that are true in any specific **M**. Axioms and their logical consequences are true in classes of models, not only in single structures. The laws of logic are universally true. Hence, the universality of truth is something different than truth-absoluteness. In the case of **AR** the class of models is narrower than in the case of logic. Yet one might remark that logical and mathematical examples cannot be extended for all truths. Although I agree that we should be careful about using the formal sciences as the philosopher's stone, one lesson that emerges from **AR** is instructive. As I remarked in Chap. 8, Sect. 8.7, mathematicians have no doubt what as to what the True Arithmetic is, and that it is identical with the theory of the standard model of **AR**. It seems that there is a strong inclination to take the standard model as *the* reality, *the* world, etc., and to say that the truths about what is standard are to be regarded as absolute (see Przełęcki 1974) regardless of whether they belong to mathematics or not. **STT** treats all models equally and has no resources to qualify one model as standard, but some other as non-standard. Once again, according to the proposed distinction, truths in all models of a given theory **Th** (it can be collection of ordinary sentences or a result of strategic considerations of the type "as if" coded by counterfactual conditionals) are true in the strong absolute sense, but truths in some models (not all) are true in the weak absolute sense. The main thesis of absolutism is that truth in a model is stable. If we take into account various species of absolutism, we do not need to distinguish proper and improper relativism, because the latter is a kind of absolutism. I guess that Kokoszyńska introduced the distinction of proper and improper relativism in order to defend **STT** against accusations of proper relativism.►

9.6 STT and Science

Some authors (see O'Connor 1975, Haack 1978, p. 113) argue that **STT** does not distinguish logical and empirical truth, because satisfaction by all sequences applies to both kinds of true sentences. As I noted in Sect. 9.2, this view confuses two separate things, namely sequences and facts. But another problem seems much more serious. If we define truth relative to **M** and **L**, the question arises as to how the concept of a model could be applied to empirical sentences, for instance in physical theories. This question was discussed by many authors—Marian Przełęcki and Patrick Suppes among others. Let **Th** be an empirical theory. According to Przełęcki (see Przełęcki 1969, Przełęcki 1974a), **Th** is a set of sentences, and its elements are true (or false) in models. Roughly speaking, Przełęcki applied **STT** directly. Suppes (see Suppes 2002) accepted the so-called non-statement view of theories, according to which they are set-theoretical predicates. A theory **Th** is true due to a suitable theorem of representation—for instance, that a measurement scale correctly represents a given class of empirical phenomena. Without entering into details, we see that the concept of a model is essential for both of the adduced views.

Even if we agree that a semantics for empirical theories is possible and, if so, it has to use the concept of model in the metalogical sense, the main problem mentioned awaits explanation. We say that empirical theories apply to empirical phenomena and, roughly speaking, their truth or falsity depends on their relation to the actual world. Yet this world is not a model in the sense of **STT**, because it is not an algebraic structure (see Wójcicki 1979, p. 157; a similar view is endorsed by other Polish philosophers of science, notably Adam Grobler and Elżbieta Kałuszyńska). Hence, one can have doubts as to whether the concept of model as elaborated for mathematical theories has any legitimate application to empirical ones. Moreover, empirical truth is approximate, but mathematical truth—unconditional, so to speak. Hence, we encounter proposals (see Wójcicki 1979, Da Costa, French 2003) which recommend the idea of approximate or partial truths that might be incorporated into logical semantics suitable for analysis of empirical theories. Another direction to the same effect consists in using the notion of verisimilitude (truthlikeness; see Niiniluoto 1987). I shall defend the position that **STT** can be directly applied to empirical sentences without any appeal to partial truth, etc. (see Woleński 2017).

Let **Th** be an empirical theory, $\mathbf{M}^{\mathbf{Th}}$—its semantic model (for simplicity, I consider a single model), and **W**—the real world. Thus, we have the ordered triple <**Th**, $\mathbf{M}^{\mathbf{Th}}$, **W**>. Since $\mathbf{Th} \subseteq \mathbf{VER}(\mathbf{M}^{\mathbf{Th}})$, the simplest solution is to assume that $\mathbf{M}^{\mathbf{Th}} = \mathbf{W}$ and say that, due to this equality, $\mathbf{Th} \subseteq \mathbf{VER}(\mathbf{W})$. However, this route is blocked for we do not know what the sign = means in this context. Inspecting the situation, we see that the role of $\mathbf{M}^{\mathbf{Th}}$ is twofold. It functions, on the one hand, as the semantic model of **Th**, but on the other, also as a model (in a rather vague r sense) of **W**. In fact, that $\mathbf{M}^{\mathbf{Th}}$ is the semantic model of **Th** does not entail that $\mathbf{M}^{\mathbf{Th}}$ is a semantic model of **W**, whatever 'being a semantic model of the real world' might mean. Assume that $\mathbf{M}^{\mathbf{Th}}$ is in some sense a model of **W**. This assumption also does not imply that **W** is a semantic model of **Th**. The above assertions suggest that something is lacking in the analysis of <**Th**, $\mathbf{M}^{\mathbf{Th}}$, **W**> (see Gårding 1977, p. 10). Possibly we can say that $\mathbf{M}^{\mathbf{Th}}$ is a mediating model between **Th** and **W** that plays a double role: as the semantic model of **Th**, and as a representation of **W** (this is suggested by Suppes' approach). In order to clarify this suggestion, we assume that we have two mappings, say, **f** and **f′**—the first, from **Th** to $\mathbf{M}^{\mathbf{Th}}$, the second, from $\mathbf{M}^{\mathbf{Th}}$ to **W**. The former is rooted in a specified interpretation $\mathfrak{I}$, the second maps $\mathbf{M}^{\mathbf{Th}}$ onto elements (empirical phenomena) in **W**. Schematically, $\mathbf{f}{:}\mathbf{Th} \longrightarrow \mathbf{M}^{\mathbf{Th}}$, but $\mathbf{f}'{:}\ \mathbf{M}^{\mathbf{Th}} \longrightarrow \mathbf{W}$. Now, in order to show how **Th** reaches the real world, there must be a third mapping, namely the composition of the mappings **f** and **f′**, defined in the standard way, that is, $\mathbf{f}'' = \mathbf{f} \cdot \mathbf{f}' = \mathbf{f}'(\mathbf{f})$, which maps **Th** into **W** via $\mathbf{M}^{\mathbf{Th}}$ as mediator (see Morgan, Morrison on models as mediators, and Magnani, Bertolotti 2017, Bueno, French 2018—on general problems related to mathematical models and their applications). **SDT** has nothing to do with the way of how the mappings **f′** and **f″** are constructed and checked. However, given **f**, **f′** and **f″**, we have

(50) '$P(a)$' is empirically true in **W** if and only if $\mathbf{v}(\mathbf{f}'')$ satisfies the formula $P(x)$.

The concept of empirical truth can be recursively defined in the standard way (see Chap. 8, Sect. 8.2). Regardless the reduction of the triple <**Th**, $\mathbf{M}^{\mathbf{Th}}$, **W**> to the pair <**Th**, $\mathbf{M}^{\mathbf{Th}}$> or the pair <$\mathbf{M}^{\mathbf{Th}}$, **W**> is not justified.

(**DG11**) The above analysis of empirical truth belongs to philosophy. Ordinary communication ignores semantic models, because linguistic devices are understood as having direct reference to the world. Scientific practice also pays no attention to semantics, because it usually focuses on models as pictorial or theoretical representations of **W**. However, if we look at what is going on in abstract physical theories, they do not deal with the real world, but its mathematical (or other—for instance, mechanical) models. Thus, if one says that theories pertain to **W**, this assumes that there is a relation or mapping between models and reality. Since if we have a mathematical model of elementary particles, for example, it is also possible to construct its semantic counterpart. That is not essential for working scientists, because the conversion of mathematical models into semantic ones does not contribute to the various procedures of creating theories or testing them. (50) and any other similar construction have its place in the philosophy of science and should be evaluated from its point of view. Tarski himself (see Tarski 1944, pp. 688–692) explicitly stressed the applicability of semantics in the methodology of science, and had quite serious expectations regarding its results. Even if such hopes should be somehow tempered today, philosophical analysis of science can profit from semantics. Yet some authors (see Schröter 1996, p. 79) prefer to speak about truth in science as provability.►

(**DG12**) How **STT** is applicable in the field mathematics? The answer depends on a general philosophical position in the philosophy of mathematics. If one says, like Girard (see Sect. 9.1), that mathematicians should not bother with the problem of truth, the issue is closed—just as with the claim that the concept of proof suffices for evaluating of correctness of mathematical statements. Yet most mathematicians maintain that if the concept of truth is legitimate in mathematics, **STT** is regarded as the best solution or quite satisfactory at least (see next digression). See Krajewski 1994 on **SDT** in mathematics, and Hales, Olivetti 1998 for a general survey of mathematical aletheiology).►

(**DG13**) Paul Benacerraf (see Benacerraf 1973, Benacerraf 1996, Benacerraf 1998, and Morton, Stich 1996, Pataut 2016 for a discussion on Banacerraf's views in the philosophy of mathematics) formulated an argument against the applicability of **STT** in mathematics. The so-called Benacerraf Dilemma runs as follows:

(a) The concept of truth is the same for mathematics and empirical science;
(b) **STT** is an explication of the traditional correspondence theory; in particular, the concept of truth is to be explained by referential concepts, for instance, via the notion of satisfaction;
(c) We accept the causal theory of perception, that is, the view that knowledge consists in causal relations between objects of knowledge and the knowing subject.

According to Benacerraf, the assumptions (a)–(c) cannot be jointly accepted for mathematical objects, because they have no causal powers. However, the premise (b) requires further explanations. First of all, it is unclear which version of **CTT** should be taken into account. It seems that the dilemma applies to the strong version, because if truth directly maps the mathematical reality (I omit the problem of empirical reality), one can argue that there must be a nexus between mathematical objects and results of mathematical knowledge. On the other hand, the weak version does need to assume such a relation. Thus, (c) is not obvious under the assumption that **STT** is based on the concept of semantic correspondence. Needless to note, that my remarks on the Benacerraf Dilemma do not pretend to explain the possibility of mathematical knowledge.►

(**DG14)** The distinction between $\mathbf{M}^{\mathbf{Th}}$ and **W** can be compared with the distinction between formal and material object of proposition (see Ingarden 1925). The latter pair of concepts (also applicable to sentences as equipped with meaning) goes back to Middle Ages. Using Ingarden's explanation, the content of a proposition determines its formal object. Roughly speaking, using contemporary terminology and omitting some puzzling examples, like negative existential, if $P(a)$ (it is enough to consider the simplest case) is a proposition, the structure **<a, P>** is the formal object of this proposition (sentence)—Ingarden speaks on the intentional state of affairs. On the other hand, the fact (state of affairs, etc.) that **a** is **P**, is the material object of the sentence in question; this entity belongs to the real world. The sentence (a) 'Zeus is the king of Olympic gods' has the formal object, but, assuming that (a) is a part of mythology, it possesses no material object. The sentence (b) 'Warsaw is the capital of Poland' has both objects. As far as the issue concerns truth, (a) is true in its formal object, but not in a material object, but (b) is true in its formal object as well as material object. The truth of (b) in the real world assumes that both objects of it agree in a sense. Clearly, there arise many additional and difficult problems, for example, the nature of agreement of both objects of propositions or connected with falsehood of mythological sentences, but I do not enter into a further discussion.►

9.7 STT and Minimalism

In Chap. 3, Sect. 3.8, I quoted few definitions of truth (see (iv), (v), (vi), (ix), (x) and (xi) called deflationary, minimalist, etc. Let 'minimalism' be the generic term for all those the related proposals to define truth (see Armour-Garb, Beall 2005, Rami 2009, Butler 2017, Cieśliński 2017), for a general presentation of minimalism and comparisons with **STT**; Paul Horwich (see Horwich 1998, Horwich 2005, Horwich 2010) represents deflationism, perhaps the most popular kind of minimalist approach to truth (see also Båve 2006). Truth-minimalism is the view that everything what is important for the concept of truth is captured by (**TS**), and is contrasted with substantivism (see Chap. 4, Sect. 4.4). Due to the role of (**TS**) in **STT**, its comparison with minimalism is natural. In particular, one can ask whether **SDT** belongs to minimalism or substantivism. I shall argue that minimalism encounters serious difficulties relative to the definition of being false and to some results from metamathematics, which cannot be accommodated within this approach to truth. Since **STT** successfully deals with both issues, the attempts to reduce it to minimalism are unwarranted.

The problem of how to define 'is false' is important for any theory of truth. It even constitutes one of Russell's conditions for a satisfactory truth-theory (see Chap. 3, Sect. 3.3). Paul Horwich agrees that the proper formulation of minimalism requires the incorporation of a definition of falsity to this approach to truth. His way out is to argue (Horwich 1998, pp. 71–73) that we cannot use the formula (the use of naïve (**TS**) suffices).

(51) A is false if and only if A is not true,

to explain 'it is false', because it would be circular even if we accept the standard truth-table for negation, that is, the equalities (i) $\mathbf{v}(\neg A) = \mathbf{1}$, if $\mathbf{v}(A) = \mathbf{0}$, (ii) $\mathbf{v}(\neg A) = \mathbf{0}$, if $\mathbf{v}(A) = \mathbf{1}$. On the other hand, (51) can be replaced by

(52) A is false if and only if not (A is true) (equivalently: A is false if and only if not-A).

If, Horwich continues, we accept the principle

(53) For any A, A is true or A is false,

we obtain (i′) $A \Rightarrow \neg\neg A$; (ii′) $\neg A \Rightarrow \neg A$; and (iii) $A \vee \neg A$. These formulas define the negation without circularity, but not completely. Horwich says (p. 72):

> A complete account of the meaning of '<u>not</u>' must contain those fundamental facts about its use that suffice to explain our entire employment of the term. Such basic regularities of use might well include acceptance of the theorems of deductive logic – which include the laws implicit in (N*) [(i′), (ii′) and (iii) – J. W.]. But a further pattern of usage, not implied by (N*), must be recognized, namely that which is characterized by the principle
>
> (K) '<u>not</u> p' is acceptable to the degree that 'p' is unacceptable.
>
> Perhaps the combination of (N*) and (K), when conjoined with the fact about the use of these terms, will be capable of explaining all our ways of deploying '<u>not</u>'. If so, then the meaning will be fixed and we proceed to define falsity in terms of <u>it</u> by means of the definition (2*) [that is, (51) – J. W.] and without fear of circularity.

I shall argue that this account is still essentially incomplete. Let us start with (i′), (ii′) and (iii′), which are tautologies of **PC**. Hence, we can say that the negation is defined inside this system. On this account, 'it is true that' becomes a truth-functional connective, more specifically, one defined by the well-known truth-tables displayed by (i) and (ii). This is a good syntactical definition, but it leaves the symbols **1** and **0** entirely unexplained. Moreover, this definition operates solely on the level of propositional language and cannot be applied to languages based on first-order logic. Hence, most applications of the concept of truth (or satisfaction) in advanced first-order semantics are lost. In particular, one cannot define tautologies of first-order logic as true in all models, etc. Adding (K) does not bring any progress in this respect. Moreover, (K) creates its own special problems. Assume that A and not-A are acceptable to the same degree. For example, one can think about tomorrow's state of weather in this way. In such a situation, (K) is incorrect. This suggests that adding (K) makes more difficulties than advantages. Summing up: the minimalist account of 'it is false' is either circular, or incomplete, or obscure. Thus, it seems that minimalism has no other way to understand falsity than its commonsense usage, but this outcome is too weak to be satisfactory for philosophical analysis.

The critical evaluation of aletheiological minimalism expressed in the last paragraph can be strengthened by the following observations. Consider two **T**-sentences: (a) 'Cracow is an old city' is true if and only Cracow is an old city, and

(b) 'Cracow is the present capital of Poland' is true if and only is the present capital of Poland. We know that (c) 'Cracow is an old city' is a true sentence and that (d) 'Cracow is the present capital of Poland' is a false one. Clearly, minimalism does not explain the difference in question (note that the issue is not how to check that the former sentence is true, but the latter false). On the other hand, **STT** actually provides an explanation. It says that (c) is true in the model corresponding to the real world, but (d)—in a different model (for example, one that refers to Poland in the fifteenth century). Although **SDT** and the minimalist definition of truth agree about (**TS**) generating concrete **T**-equivalences, the former explains why (c) is true, and (d)—false, but the latter stops at (a) and (b)—both true. Yet (a) consist of two truths, but (b)—of two falsehoods. Similarly, in the case of future contingences, A and $\neg A$, we have two true **T**-sentences but nothing more can be said. Yet if the minimalist were to say "Well, I exchange classical logic for three-valued logic", such an answer would lead to further difficulties for the lack of the minimalist's approach to many-valuedness. The arguments advanced in this paragraph provide additional reasons to regard minimalism as an incomplete theory of truth.

Tarski offered the following argument against the so-called nihilist theory of truth (Tarski 1944, p. 683). Consider the sentence

(54) Consequences of true sentences are true.

According to nihilism (it is a species of minimalism; Tarski was influenced by Kotabiński's distinction of the real and verbal use of 'is true' (see Chap. 3, Sect. 3.7.6), the adjective 'true' can be deleted from every context in which it occurs without any loss of content. Clearly, we cannot erase 'true' from (54). It does not preclude (54) from being transformed into a sentence which covers it content. One might, for instance, propose

(54) For any A, B (if A and $A \vdash B$, then B).

Then, by using (**TS**), we obtain

(55) For any A, B (if A is true and $A \vdash B$, then B is true).

Since the provability relation is syntactic, it does not require semantics. However, we still need to justify the use of quantifiers, and that it is not easy without appealing to semantics.

Even if we accept (6) as a correct paraphrase of (4), minimalism has very serious difficulties with metamathemical results. The main problem concerns limitative theorems (see Chap. 8, Sects. 8.4–8.5). Some authors (see Ketland 1999, Glanzberg 2003, Ketland 2005) argue that minimalism does not have sufficient resources to express the undecidability of **AR** (and, a fortiori, the undefinability of truth), but others (for instance, Tennant 2002, Tennant 2005) try to defend the minimalist approach to Gödelian phenomena (this phrase is frequently used in relevant debates). It seems that all syntactic devices—arithmetization and diagonalization, in particular—are accessible to "minimalistic" metamathematics. As far as Gödel's informal argument for undecidability (see Chap. 8, Sect. 8.3) is concerned, the minimalist can say that it is just a proviso, but that the official proof has a purely syntactic character. Real difficulties for minimalism appear when semantics enters into metamathematics. At first sight, the minimalist might say that the **T-**predicate for **AR** is definable if all **T**-equivalences are derivable in **AR**. However, this move does not suffice, because the extension of the truth-predicate is not defined by the collection of its arguments; otherwise would be inconsistent. This assertion follows from (**FPL**). Now, if the minimalist will say that he or she uses the predicate 'is true' in the intuitive sense, this is a way of saying that minimalism is incomplete. Moreover, and perhaps more importantly, some model-theoretic metamathematical results are not expressible in the minimalist language—for instance, the theorem that models of **AR** are not definable in $\mathbf{L_{AR}}$. Neil Tennant argues that the Löb theorem (if for any A, **AR** $\vdash$ 'A' is provable in **AR**, then A if and only if **AR** $\vdash A$), that is, the so-called reflection principle (roughly speaking, what is provable, is true) helps the minimalist to account for Gödelian phenomena. However, this argument has three weak points, namely (i) Löb's theorem assumes that **AR** is consistent, but consistency is not provable in **AR**; (ii) not all arithmetical truths are provable in **AR** (see (i)); (iii) minimalism is too weak to distinguish standard and non-standard models.

The above considerations show that minimalism is too weak in some important respects. Firstly, its account of false sentences has serious limitations since **T**-sentences do not suffice for clarifying of the distinction between being true and being false (**T**-scheme holds for falsehood as well). Secondly, **T**-sentences fail in the case of future contingencies in this sense that **SDT** does not entail contradictory **T**-biconditionals. Thirdly, minimalism cannot accommodate many model-theoretic results in metamathematics. Consequently, a stronger truth-theory (it must be just substantive) is required not only for philosophy, but also for metalogical investigations. That philosophers need a stronger theory of truth than that offered by minimalism is an important point, because one could claim that cognitive interests of logic and mathematics have no particular significance for the philosophical enterprise. The fate of minimalism shows that its inaccuracies do not only involve the formal sciences, but also philosophical tradition. Even if **STT** were qualified as a wrong approach to the concept of truth, its strength far exceeds what minimalism has to offer. Be that as it may, **STT** cannot be reduced to the minimalist theory of truth.

(**DG15**) To complete terminological remarks on aletheiological deflationism, let me note that its opposite, that is, substantivism, is sometimes termed as inflationism in the theory of truth; hence, the correspondence theory is regarded as inflationary. Due to negative associations with economic facts related to circulation of money, I have reservations concerning this terminology. On the other hand, the term 'deflationism' is so popular nowadays, that it would be hard to resign from its usage. In fact, applications of the concept of deflation I economy as less known than of inflation.►

9.8 STT and Coherentism

In Chap. 3, Sect. 3.3 and Chap. 4, Sects. 4.4–4.5 I mentioned that the coherence theory of truth has two versions: Neo-Hegelian (Bradley and his followers) and positivistic (this is an ad hoc label). The former is based on a very special logic (see below), but the latter assumes classical logic (for a general analysis of coherence theories, see Khatchadourian 1961, Rescher 1973, Walker 1989, and Olsson 2005). In his criticism of the coherence theory Russell argued (see, for instance, Russell 1984, p. 150) that beliefs in dreams could be regarded as coherent "in any sense of the word which seems admissible", and, thereby true. Assume that Russell was thinking about coherence as consistency. However, no coherence theory reduces truth to consistency. For Bradley, coherence as truth requires comprehensiveness, but Neurath required sensitivity with respect to empirical data.

Due to the above explanations, we have

(56) $\vdash (\mathbf{X} \in \mathbf{COH} \Leftrightarrow \mathbf{X} \in \mathbf{CONS} \wedge \mathbf{X} \in \mathbf{Ch})$,

where the symbol **Ch** refers to an additional condition of coherence. Due to this statement (the sign $\vdash$ is admissible, because the coherentist regards (56) as a theorem), consistency acts as a necessary condition of coherence, but not as sufficient. Thus, we have

(57) $\vdash (\mathbf{X} \in \mathbf{COH} \Rightarrow \mathbf{X} \in \mathbf{CONS})$.

By monotonicity of provability, we obtain

(58) $\vdash (\mathbf{X} \in \mathbf{COH}) \Rightarrow \vdash (\mathbf{X} \in \mathbf{CONS})$.

By contraposition, (58) is transformed into

(59) $\neg \vdash (\mathbf{X} \in \mathbf{CONS}) \Rightarrow \neg \vdash (\mathbf{X} \in \mathbf{COH})$.

The last formula means that the unprovability of consistency entails the unprovability of coherence. Thus, if the coherence theory of truth is intended to define truth for **AR** and its extensions, (59) shows that this task cannot be fulfilled. Even an informal version of (**TG1**) justifies this assertion. Hence something must be added to consistency in defining coherence, and the condition **Ch** plays this the role; all known proposals that aim in this direction are not sufficiently precise.

Coherence can also be considered from the model-theoretic (semantic) point of view via the completeness theorem in the version (see Chap. 5(26)) that a set of sentences **X** is consistent if and only if it has a model. In order to satisfy the comprehensiveness condition, we can even add that **X** is maximally consistent. At any rate, these paraphrases fit to some extent Bradley's intuitions that as the body of truths **X** and its model are infinite realities. I say 'to some extent', because there is one Absolute (= Reality) for Bradley. We can be still closer to Bradley, if the standard model is taken into account. One could say that **X** is maximally coherent and the Absolute constitutes its model. The coherence in question can be qualified as semantic. Assume that **AR** is a part of **X**. Consequently, proving that **X** is complete requires infinitary methods accessible in the metatheory of **X**. And now a new problem arises. One of the motives for coherentism (see Chap. 4, Sect. 4.5) is the lack of truth-criteria in (**CTT**). Moreover, such criteria should afford a definite decision as to whether a given sentence is true or not. If so, the proof of the semantic coherence of **X** should be effective, that is, performable in a finite number of steps. For instance, Nicholas Rescher (see Rescher 1985, p. 797) says that if **CrT** is to be an adequate truth-criterion, then the truth of any sentence must consists in the satisfaction of **CrT** by this sentence. However, if we agree that **AR** defines the upper limit of effective methods, the coherentism claim cannot be realized in a manner acceptable to representatives of the discussed approach to the concept of truth. In particular, consistency, a necessary condition of coherence, is not effectively provable in the case of sets of sentences containing **AR** (see (59) above). Thus, coherentism land on the following horn: if **CrT** is effective, it does not satisfy itself; but if it can be non-effective, the semantic coherence winds up being the same as **SDT**. This conclusion holds for Neo-Hegalianism as well as for positivism (in the above provisional meaning).

Bradley and his followers could defend his approach by saying that concrete sentences are partially true, but the full (entire) truth can be predicated only about the maximally consistent set. This argument is at odds with metalogic. Every true set of sentences is compact (see (**Df15j**) in Chap. 5) since it is consistent. This means that if **X** is an infinite set of true sentences, all its finite subsets are also consistent. However, the Bradleyan truth is not compact. Although **X** (assume that it represents the Bradleyan truth) is true in the absolute sense, its fragments are not true in the same sense. And conversely, although finite subsets of **X** are partially true, the Bradleyan truth is not, because it contains "the entire truth and only truth". Thus, Bradley uses two different notions of truth without a clear explanation of the relation between them. Clearly, any finite set of partial truths can be (it remains question whether it is, but I leave this problem without further comments) internally consistent, but its coherence in Bradley's sense is very problematic. Perhaps this circumstance provoked Russell to the following remark (see Russell 1984, p. 149):

> *The coherence-theory* is generally advocated in – for example in Mr. Joachim's *Nature of Truth* – in connection with logic wholly different ours. The chief arguments against it are arguments against the logic with which it is commonly associated; but we shall assume these arguments and confine ourselves to a statement and a refutation which assume our logic.

Perhaps my considerations on the Bradleyan truth and compactness provide good evidence for Russell's point expressed in the above quotation. Anyway, an obvious advantage of **STT** over the Bradleyan coherence theory consists in the former assuming a well-known logic, whereas the second fails in this respect. As far as the positivistic coherentism is concerned, it is essentially incomplete due to (59). In particular, the unprovability of coherence cannot be improved by the satisfaction of additional constraints that are imposed on empirical criteria of acceptance. Although logic used by Neurath or Carl Hempel (the main representatives of the positivistic coherentism) was "normal" (in fact, classical), the set of true sentences in a model **M** cannot be determined by **Ch**. In other words, there are truths undecided by the coherence criterion.

9.9 STT and Realism

'Realism' is one of the most ambiguous words in philosophy. It is immediately clear when we consider its denotations as determined by various adjectives. Here are some examples: epistemological, ontological, metaphysical, semantic, direct, naive, critical, scientific, common-sense, internal, transcendental, global or practical. Other instructive hints are associated with various oppositions, like realism/nominalism, realism/idealism, realism/anti-realism, realism/phenomenalism, etc. Is there any common core to all species of realism? Perhaps the most general is that realism with respect to objects of a kind **K** consists in admitting their existence independently of the existence of objects of another kind **K′**. Thus, we are realists concerning universals (the opposition of realism to nominalism) if we say that they exist independently of individuals. We are realists concerning the objects of knowledge if we say that we cognize something which is independent of our epistemic acts. I shall not multiply the examples. What I said is sufficiently to show how complicated and multidimensional is the issue of realism in philosophy and its particular branches, not only ontology and epistemology, but also ethics or aesthetics.

(**DG16)** According to Tarski (Tarski 1944, p. 686):

> [...] we may accept the semantic conception of truth without giving up any epistemological attitude we may have had; we may remain naive realists, critical realists or idealists, empiricists or metaphysicians – whatever we were before. The semantic conception of truth is completely neutral toward all these issues.

I agree with this view of Tarski insofar as the issue of logical entailment between realism (in any meaning) and **STT** is concerned. This means that **STT** as a logical construction does not imply any kind of realism and is not implied by any form of realism. We get a different perspective if we consider **STT** as a piece of philosophy. Recalling my philosophical methodology (see Introduction), I regard logical constructions as a source of philosophical insights via suitable paraphrases, at least in analytic philosophy. Thus, I shall argue for semantic realism (**SR** for brevity) and try to show its consequences for epistemological realism (**ER**). Its oppositions are: semantic anti-realism (**SA**) and epistemological anti-realism (**EA**). Finally, I shall note some ontological problems related to **ER** and **EA**. This section is based on Woleński 2004. Note that I limit my remarks only to **STT**. The topic "Truth and Realism" is much more extensive (see Devitt 1996, Gardiner 2000, Greenough, Lynch 2006, Taylor 2006, and Novák, Simonyi 2011).►

I begin with a characterization of **SR** and **SA** in McGinn 1980 (see also Hinzen 1998 for a more comprehensive analysis). Colin McGinn introduces the following scheme to capture a particular form of this controversy ('realism' means 'semantic realism' and 'anti-realism'—'semantic anti-realism'):

(60) realism $\Rightarrow$ (meaning $>$ use),

where ‘meaning’ stands for ‘meaning of a sentence’, ‘use’—for ‘use of a sentence’, and the symbol > means ‘transcends’. If we now assume that ‘x does not transcend y’ means ‘x is equal to y’, then by contraposition of (60) we obtain

(61) (meaning = use) $\Rightarrow$ anti-realism.

However, (60) and (61) should be strengthened to equivalences in order to achieve a full comparison of realism and anti-realism. Thus, we can propose,

(62) realism $\Leftrightarrow$ (meaning > use).
(63) anti-realism $\Leftrightarrow$ (meaning = use)

as a preliminary characterization of the standpoints under discussion. Normally, realism is supplemented by the truth-conditional theory of meaning (meaning is defined by truth-conditions) (see Davidson 1967 and Sect. 9.4 above), whereas anti-realism replaces use by assertibility-conditions (see Sect. 9.8). This leads to

(64) realism $\Leftrightarrow$ ((meaning = **TrC**) and (**TrC** > **AsC**);
(65) anti-realism $\Leftrightarrow$ (meaning = **AsC**).

where **TrC** = truth conditions, **AsC** = assertibility conditions. Now the question arises as to whether ‘meaning’ in (6) is replaceable by ‘truth-conditions’. Any answer depends on how truth-conditions and assertibility–conditions are understood.

According to Davidson, truth-conditions are generated by (**TS**). This means, however, that truth-conditions are not generally given in a constructive (effective) way, because **STT** is not effective. On the other hand, the anti-realist claims that the constructivity of assertibility-conditions imposes a basic constraint on **SA**. If we drop this constraint, both realism and anti-realism would be practically indistinguishable. To see this, it is sufficient to consider

(66) A is assertible if and only if it is possible to describe its truth-conditions.

If we adopt (66), then the difference between **TrC** and **AsC** simply disappears. Therefore, (64) and (65) wind up saying the same. This shows that the equality

(67) **AsC** = constructive **AsC**

is actually important for the realism/anti-realism issue. Moreover, constructive **AsC** have their source in a constructive logic.

An appeal to the semantic definition of truth plays an essential role in obtaining (67) because it explains why **STT** generates truth-conditions non-effectively, due to (**BI**). Instead (64) and (65), we can propose

(68) realism $\Leftrightarrow$ (truth > **AsC**),
(69) anti-realism $\Leftrightarrow$ (truth = **AsC**).

The last two formulas show explicitly that the realism/anti-realism controversy focuses on how the concept of truth should be defined. The way that leads to (68) suggests that any definition of truth consistent with this reasoning must imply bivalence. Hence, if we replace ‘truth’ in (68) by ‘truth under **STT**’, a particular form of realism is achieved; this is perhaps a possible reason why Tarski's conception of truth is usually realistically interpreted, perhaps even against Tarski himself (see (**DG10**). Anyway, the term 'truth' cannot be understood in the same way in both (68) and (69) for different meanings of 'truth-conditions' For the anti-realist, the last term means something like '*A* is verifiable', '*A* is constructively asserted', '*A* is provable', etc. Such explanations force rejecting up (**BI**), which is forced by **STT**. Assume now that 'truth' in (69) is replaced by 'truth under **STT**'. Clearly, the assertibility conditions cannot be understood in the anti-realistic manner. Even if the anti-realist says that (**TS**) can be incorporated into the anti-realistic semantics, that requires additional stipulations (see Sect. 9.8 above).

I consider semantic realism to be sound. However, we need an *experimentum crucis* in order to argue that the realism/anti-realism controversy should be solved in favour of realism. It seems that (**TG1**) and (**TG2**) are useful in this respect. More specifically, the existence of Gödelian sentences suggests that anti-realism is wrong. Anti-realists argue that we understand sentences if we can effectively describe their assertibility-conditions. Although the existence of unprovable does not force the a priori rejection of anti-realism, at least one point in anti-realism so construed still remains unclear. It is captured by the question of how the anti-realist can assert that he or she could grasp the meaning of an undecidable sentences without knowing whether it has a constructive proof. It seems that the main problem for anti-realists stems from the existence of undecidable (not only undecided) problems, because we should clearly distinguish between the two following statements:

(70) It is not known whether *A* is constructively provable;
(71) It is known that *A* is not constructively provable.

Assume (see Sect. 9.8 above) that a proof is constructive (effective) if and only if it is formalizable within **AR**. Under this convention, Gödelian sentences have no constructive proof. Moreover, since **AR** is interpretable in **HA** (the Heyting arithmetic (see Chap. 8, Sect. 8.3.1), the latter is incomplete and its consistency is unprovable in it. Thus, Gödelian sentences exist also in intutionistic mathematics, and if anti-realist semantic principles are to be considered seriously, Gödelian sentences have no meaning and are not intelligible. By adopting a special convention concerning the concept of meaning, we might perhaps argue that Gödelian sentences have no meaning. On the other hand, the statement that they are not graspable is certainly empirically false. For realism, Gödelian sentences are true or false, and, a fortiori; mea-ningful and understandable. This is a difference, perhaps even the main one, between realism and anti-realism. If we look closely at (70), it becomes clear that in order to know that *A* is constructively provable we must already understand it. The only way out for the anti-realist is to say that (70) has no significance for him and stay with (71) only. However, that is a very weak position. Having (71), the anti-realist can only say that he is uncertain about the meaning of *A* which is known as constructively not provable, but that is a typical case of begging of question.

By virtue of (**TG2**) the statement

(72) Mathematics is consistent,

belongs to the class of Gödelian sentences and is not effectively provable, even though it does have a model-theoretic justification. If we construct a model for a given system, we also have an argument for its consistency—even if possessing of this property it is not provable effective means. Is (72) important for the anti-realist? Absolutely—because the anti-realist appeals to (72) as a criterion for the correctness of mathematics. Though the anti-realist does not identify existence with consistency, the latter is a necessary and sufficient condition of existence. Yet if the anti-realist who uses (72) is not able to prove it effectively, he or she either does not know its meaning, must essentially modify his semantic principles. What is the moral to be derived from Gödel's theorems for the discussed issue? The moral is exactly that truth transcends assertibility-conditions. It gives a strong philosophical argument for (68)—the main thesis of semantic realism.

As far as the issue the relation of **SR** and **SA** to **ER** and **EA** is concerned, (**TT**) can help us to illuminate this question. I will restate this theorem as

(73) If **Th** is a theory, **AR** $\subseteq$ **Th**, and $\mathbf{M}^{\mathbf{Th}}$ a model of **Th**, the expressive power needed for defining $\mathbf{M}^{\mathbf{Th}}$ > the expressive power of **Th**.

A straightforward epistemological interpretation of (73) is surely realistic. We can identify the expressive power of a conceptual apparatus with something that is essential for the Knowing Subject. Thus, if we say that the expressive power needed for the definition of **M** transcends the expressive power accessible for the Knowing Subject, we thereby suggest that **M** transcends the Subject in some respect. This figurative description can be replaced by interpreting **M** as the object of knowledge and **Th** as expressed in the object language used by the Knowing Subject, we have:

(74) **ER** $\Leftrightarrow$ (the expressive power needed to define of the object of knowledge > the expressive power of **Th**);
(75) **EA** $\Leftrightarrow$ (the expressive power needed to define the object of knowledge = the expressive power of **Th**).

Speaking loosely, **SR** maintains that the object of cognition transcends the conceptual apparatus (represented by **Th**) possessed by the Knowing Subject. In other words, there is always a residuum of **M** which exceeds **Th**. Enriching **Th** by new expressive means does not change the situation, because the new model **M′** is not fully captured by **Th′** as an extension of **Th**. On the other hand, according to **EA**, **M** does not transcend the conceptual apparatus of the Knowing Subject.

One could still remark that the Knowing Subject uses not only **L** but also **ML**, and that the latter language enriches the conceptual apparatus used in speaking about the object of knowledge in a way entirely internal (subjective). In other words, the object of knowledge is constituted as construction by the subject (the thesis of epistemological idealism). However, the description of the cognitive situation in **ML** inevitably involves the relation between theory **Th** formulated in **L** and the model **M**. The main thesis of **ET** can be expressed as

(76) **M** exists independently of **Th**, because it cannot be defined by the expressive means of this theory.

The realist is inclined to read (76) ontologically. The denial of (76) can be achieved by adding the Berkeleyan dictum *esse* = *percipi*, or perhaps the assertion that to be = to be completely described by **Th**—this moves produces anti-realism as subjective idealism. Scepticism (there is nothing ontological to be derived from (76) or its denial), Kantianism (*Dinge an sich* exist, but are unknowable) or Putnam's internalism (models are internal entities), are there other possible ways out. Now a typical anti-realist strategy is to explain why ontological theses are not derivable from statements like (76). The anti-realist maintains that his realist opponent make use of a wrong theory of meaning, and adds that adopting the concept of meaning governed by (69), blocks inferences that support realism. Anyway, there is nothing to prevent anti-realism from adding the ontological assertion to the denial of (76), and saying that the real world exists but that statement is not derivable from any semantic considerations pertaining truth-values of sentences.

I explicitly repeat once again that I consider semantic, epistemological and ontological consequences presented in the above analysis not as logical consequences of metalogical theorems, but as interpretative outcomes obtained by some hermeneutical means applied to the readings of logic in order to get philosophical insights. We should then tell how we translate traditional philosophical standpoints into our new language, motivated by **STT** in the case of defending **ER**, for instance, what does it mean that the world exists independently of our cognitive acts. The anti-realist denies all of that. If the anti-realist uses metalogic, his or her situation is exactly parallel to that of the realist. We can take Putnam's internal realism (which, ironically, is a kind of anti-realism) as an example. He uses the Löwenheim–Skolem theorem (if a theory has an infinite model, it also has a countably infinite model; see Chap. 5(29)), and concludes that it allows us to speak about internal models only. However, this theorem does not say anything about external or internal models, but only about models *simpliciter*, that is mathematical structures, and nothing ontological follows directly from this metamathematical result without undertaking hermeneutical moves toward its philosophical understanding.

Certainly, we are not entitled to say that the traditional controversy between realism and anti-realism is solved by metalogical methods (although I follow Ajdukiewicz 1937 and Ajdukiewicz 1948, but my conclusions are more tempered; according to Beth (see Beth 1968, p. 621) considers Ajdukiewicz's criticism of idealisms as conclusive). On the other hand, we can show what is involved when we make inferences based on the premise that we possess or recommend theories of the existence of the real world. In fact, doing such reasoning we present the old problem of the relation between epistemology and ontology—already dramatically unveiled by the ancient Stoics. Or we come back to the famous problem raised by Brentano: what are the objects to which our mental acts are directed? Are they understood as parts of intentional acts, or exist as independent entities? Or we approach a more recent question: when I say 'I know that *A*', am I making *de dicto* or *de re* assertion*?* It is quite fascinating that these (and other) traditional problems have a metalogical setting, Yet, to repeat, we should be very careful in our claims. In order to illustrate that, consider the succession of views:

(77) Ontological Realism $\Rightarrow$ Epistemological Realism $\Rightarrow$ Semantic Realism,

Some philosophers consider (77) literally. However, this view must be qualified as an oddity, because it calls for rejecting ontological realism as a consequence of refuting semantic realism. That means that a view concerning semantic matters forces us to an ontological decision—which is (or at least, can be) considered very strange. It seems that relations between the views mentioned in (77) cannot be reduced to purely logical connections. It is rather the case that these views have ontological, epistemological and semantic aspects simultaneously. A philosophical virtue of **STT** consists in the fact that this account of truth provides tools for looking at interconnections between these aspects from the point of view of formal semantics.

9.10 Two Application of STT to Analysis of Knowledge

The classical definion of knowledge defines it as true justified belief. This definition looks as fairly satisfactory for the first sight. If someone, let say, a person *X* says (*) 'I know that *A*' and adds that he or she does not believe that *A* and/or has no justification for *A*, we are inclined to deny that *X* knows that *A*. The same conclusion stems from demonstrating that *A* is false. On the other hand, most scientific and commonsense assertions are subjected to revisions or even radical changes. Copernicus belief that planetary orbits are circular was rejected in the favour of the statement that they are elliptical. Yet we are not ready to say Copernicus had no knowledge about the motion of celestial bodies. Consequently, the fact that someone (a scientist, a common-sense knower) has false beliefs does not decide that he or she is lacking of knowledge. One of proposals to solve this puzzle consists in employing the idea of verisimilitude (see Sect. 9.6), but I will not discuss this solution. I will concentrate on more epistemological issues.

The Gettier problem (see Gettier 1963 and Shope 1983 for a summary of discussions in 1963–1983) is a famous example of problems related to the classical definition of knowledge. Edmund Gettier uses the following definition of knowledge:

(78) *X* knows that *A* if and only if
(a) *A* is true;
(b) *X* believes that *A*;
(c) *X* is justified in believing that *A*.

Two additional constraints are covered by

(79) (a) The condition (79c) does not preclude that *A* is false;

(b) for any *A* and B, if $A \vdash B$ and *X* is justified in believing that *A*, *X* is also justified in believing that *B*.

Assume (it is Gettier's original example) that the persons *X* and *Y*, both men, applied for a certain job. Suppose that *X* is justified in believing in the conjunction (i) '*Y* is the man who will get the job, and *Y* has ten coins in his pocket'; X's evidence is based on the standpoint, known to *X*, of the boss deciding on the competition and counting (by X) coins in the pocket of Y. The sentence (i) entails (ii) 'the person who will get the job has ten coins in his pocket'. By (79b), *X* is justified (even strongly) in believing that (ii). However, the boss changed his opinion and decided *X* gets the job. Moreover, *X* has also ten coins in his pocket, but he does not know that. Hence, (ii) is true, *X* believes that (ii), and *X* is justified in believing that (ii). Thus, according to (78), *X* knows that (ii). However, this conclusion is counter-intuitive, because *X* thinks about *Y* that the latter get the job, but this belief is false.

The second example to be discussed in this section was given by Jonathan Dancy (see Dancy 1985, p. 25). He says (it is his version as a Gettier-like argument):

> Henry is watching the television on a June afternoon. It is Wimbledon men's final day, and the television shows McEnroe beating Connors; the score is two set to none and match points to McEnroe in the third. McEnroe wins the point. Henry believes justifiably that
>
> 1. I have just seen McEnroe win this year's Wimbledon final and reasonably infers
> 2. McEnroe is this year's Wimbledon champion.
>
> Actually, however, the cameras at Wimbledon have ceased in function, and television is showing a recording of last year's match. But although it does so that is in the process of repeating last year's slaughter. So Henry's belief 2 is true, and surely he is justified in believing 2. But we would hardly allow that Henry knows 2.

Dancy's example uses implicitly (79b), because inference from 1 to 2 tacitly uses the premise that the winner of Wimbledon final is the Wimbledon champion in the given year. The conclusion "we would hardly allow that Henry knows 2" is justified by the fact that Henry did not watch the actual final, but the match one year later.

(**DG17**) Russell (see Russell 1948, p. 170/171) anticipated Dancy's example:

> It is clear that knowledge is a sub-class of true beliefs: every case of knowledge is a case of true belief, but, but not vice versa. It is very easy to give examples of true beliefs that are not knowledge. There is the man who looks at a clock which is not going, though he thinks that it is, and who happens to look at it at the moment when it is right; this man acquires a true belief as to the time of day, but cannot be said to have knowledge.

Although Russell did not explicitly mention the condition (78c), but certainly considered perceiving the state of a clock as a justification of the belief in question.►

There are many proposals how to solve Gettier's like counterexample (see surveys in Shope 1983 and Dancy 1985, pp. 27–34). Typically, they consist in adding an additional condition to (78), for instance, prohibiting inferences from false premises. I will employ the difference between formal and material object of sentences (see (**DG14**) as the starting point. All Gettier's examples regard involved trouble-making beliefs as concerning the real world and qualified erroneously as being true and sufficiently justified. In other ways, they are intended as expressed in true sentences about the real world (the condition (78c) is not relevant for my further discussion). The sentences "*Y* is the man who will get the job, and Y and has ten coins in his pocket is the man who will get the job, and *Y* has ten coins in his pocket', 'the person who will get the job has ten coins in his pocket' and 'McEnroe is this year's Wimbledon champion' has formal objects involving *Y*, his ten coins in pocket, the person (identified as *Y* by *X*) having ten coins in his pocket and McEnroe as the Wimbledon champion in the given year. We can easily construct models semantic models of these sentences as suitable mathematical objects, in particular, model-theoretic reducts (see Chap. 8, Sect. 8.2). However, truth in such reducts does not imply truth on the real world, because there is not composition of functions mapping truth in reducts to truth on the real world (see Sect. 9.6). And this situation is responsible for possible incorrectness of inferences about knowledge. For instance, if *X* thinks about the sentence 'the person who will get the job has ten coins in his pocket' as referring to himself, his conclusion is false as asserting something about the real world, but true in the semantic model of this sentence as construed by the speaker. A similar analysis of Dancy's example requires considering the sentence 'According to what I watched on TV as showing the actual final, McEnroe is the champion' which is semantically true in its semantic model (its formal object) but false on the real world, because TV did not show the actual final. **STT** in its model-theoretic version allows a precise account of what is going on in Gettier's counter-examples. Of course, scientific practice and daily life ignore the distinction of two objects of propositions (see (**DG11**)) and the related formal constructions, because in most cases formal and material objects coincide.

The above considerations suggest a modification of (78), captured by

(80) *X* knows that *A* if and only if
(a) believes that *A*;
(b) *A* is true in $\mathbf{M}^{\mathbf{A}}$;
(c) $\mathbf{M}^{\mathbf{A}}$ is congruent with **W**;
(d) *X* is justified in believing that *A*.

I am fully aware that (80c) and (80d) require further considerations. Anyway, the role of **STT** is obviously important in (80).

Suppose that we consider the knowledge in its integrity (symbolically **Knowl**), that is, all possible cognitive results. Linguistically, **KN** consists of all true sentences in a model $\mathbf{M}^{\mathbf{KN}}$ assuming that the condition (80c) is satisfied. Clearly, **KN** is infinite. Since our cognitive tools are finite, we have no chance to an effective (in finite number of steps) reaching **KN**. Is this circumstance an argument for epistemic pessimism—a view that truth is not obtainable by humans? Suszko's (see Suszko 1968) provided an interesting formal model of the development of knowledge. Let the symbol **VER**(**M**) refers to the **KN**. Due to an earlier remark this collection is out of human cognitive possibilities. What can we do consists in accumulating fragments (subsets) of **VER**(**M**), represented by sets $\mathbf{VER_1}(\mathbf{M})$, $\mathbf{VER_2}(\mathbf{M})$, $\mathbf{VER_3}(\mathbf{M})$, …, $\mathbf{VER_k}(\mathbf{M})$, … such that for any $i(1 < i < k)$, $\mathbf{VER_i}(\mathbf{M}) \subset \mathbf{VER_k}(\mathbf{M})$; consequently, for any i, $\mathbf{VER_i}(\mathbf{M})$ is a part (a sub-class) of **VER**(**M**). The definition

(81) $$\lim_{\mathbf{i} \to \infty} \mathbf{VER_i}(\mathbf{M}) = \mathbf{VER}(\mathbf{M}).$$

This equality means that **KN** is the limit of the sequence of partial set of truths (not the sets of partial truths!). Roughly speaking, every truth can be discovered, it is impossible to discover all truths altogether and every error can be corrected. This model assumes the cumulative picture of the development of knowledge. Moreover, it is simplified by supposing that **KN** is represented just by the unique model—it is quite possible that we need many models in order to cover the plurality of real facts. On the other hand, the connection of (81) with **STT** is obvious, because properties of **VER**(**M**) display the fact that the collection of all truths, including mathematics, is not definable in the language assiociated with **M**.

9.11 Conclusion

If we look at **SDT** within metalogic, it functions as a normal definition in the sense accepted in the traditional or contemporary theory of definition. Its preparatory part concerns the concept of satisfaction and is inductive—the ultimate definition has the form of equivalence. On the other hand, the status of **SDT** inside philosophy is fairly problematic and calls for further explanations. In fact, definitions in philosophy do not have a good reputation, and this situation has very convincing reasons. So-called analytic definitions have to cover so many intuitions, frequently—mutually conflicting or non-homogenic, that their explanatory effectiveness is very limited, particularly, as successful means of exhibiting meanings of relevant concepts (think on the notion of realism, for example, but many other suggestive illustrations immediately comes to mind). On the other hand, if someone proposes a synthetic or regulative definition, a natural question is why some intuition is chosen as fundamental, whereas other are neglected or considered as secondary, or why a normalized decision is made at the given point but not at different one; moreover, regulative definitions are always conventional to some extent and subjected to the objection that they ignore the essence (a beloved category of philosophers) of defined concepts or entities. Moreover, a chance for producing a short (of course, taking this property *cum grano salis*) formulation falling under the scheme '*X* is *Y*' where the letter *X* refers to a definiendum, but the letter *Y*—the definiens associated with what is defined, seems rather small. This conclusion is sufficiently confirmed by the entire history of philosophy.

Taking into account these circumstances as sufficiently indicate various difficulties in using definitions in doing philosophy (I do not suggest that philosophy is exceptional in this respect, because the similar situation occurs in the humanities as well as it is documented by various attempts to define art, religion, music, language, etc.). I guess that other analytic procedure has a greater significance for philosophical enterprise. I think here about Carnap's idea of explication (see Carnap 1950, pp. 3–8) as a much better servant for philosophical needs than formulas having the standard shape of definitions in the sense of elementary (school) logic.. The method of explication consists in using and comparing two ingredients, namely an explicandum (what is explicated in a given conceptual situation; analogy to definiendum) and an explicans (by which the explicatum is explicated—analogy to definiens; one can say that explicatum is clarified, explained, illuminated, specified, etc.). I would add at this place that interpretative consequences (see Introduction) have much more to do with explications in Carnap's sense that with definitions. In fact, I consider my investigations undertaken in the book as interpretative consequences derived (via hermeutics which I tried to exhibit in an explicit manner) from a well-established logical theory.

Truth in the ordinary meaning, or even scientific is the explicandum on which the present book focuses (some authors speak of the pre-theoretical sense of being true in this context), but **STT** (not only **SDT**) functions as the actual explicatum. Carnap formulated the following constraints (conditions of adequacy) for the procedure of explicating (they pertain to properties of explicanda):

(a) similarity to the explicandum;
(b) exactness;
(c) fruitfulness;
(d) simplicity.

Although (a)–(d) are highly evaluative (it is nothing strange in philosophy), Carnap himself believed that, they might serve as tools for transforming philosophy into science (his analysis addressed the concept of inductive probability as explicated by the mathematical axiomatic theory of probability). I have no such ambitions vis-a-vis the concept of truth, and my task aims at a philosophical clarification. On the other hand, the points (a)–(d) can be also used in evaluation of this more moderate intention. Checking whether **STT** satisfies Carnap's constraints one can observe:

Ad (a) In (**DG4IX**), I argued that **STT** is close to the ordinary usage of 'is true', even more closer than other popular philosophical theories of truth. I consider this fact as a very serious advantage of the semantic theory of truth;
Ad (b) **STT** is exact owing to its formulation in logical and metalogical terms;
Ad (c) **STT** enables us to clarify many problems of traditional aletheiology;
Ad (d) **STT** is simple—at least for those who know contemporary logic.

When I studied philosophy with Roman Ingarden, he, as a faithful phenomenologist, denied that logic could have a deeper philosophical significance. According to him, **SDT** should be interpreted as a logical construction, relevant for logic, but not for philosophy. Is **STT** logical or not? Well, Tarski's definition does not belong to pure logic (in the understanding outlined in Chap. 5), but to metalogic. I included in previous chapters a considerable amount of logic and metalogic (or metamathematics) in order to show that even if the semantic theory of truth is not regarded as the best, it has sufficient virtues to be considered as a serious philosophical doctrine. However, I feel myself obliged to say that most statements in this book, should be taken as a piece of philosophical speculation built on logic. Finally, I would like point out still one virtue of logical analysis. As I noted at the end of Introduction, Tarski is not sacrosanct. The same concerns **STT**. Although I defend this theory, I am fully aware that it is opened to criticism. The virtue in question consists in the fact that logic makes criticism easier than other philosophical devices, because the former forces transparency of reasoning and its results.

References

Achourioti, Th., Galinon, H., Fernández, J. M., Fujimoto, K. (Eds.) (2015). *Unifying the Philosophy of Truth*. Dordrecht: Springer.

Agazzi, E. (Ed.) (2017). *Varietes of Scientific Realism*. Dordrecht: Springer.

Ajdukiewicz, K. (1937). Problemat transcendentalnego idealizmu w sformułowaniu semantycznym (A semantical version of the problem of transcendental idealism). *Przegląd Filozoficzny*, 40, 271–287; Eng. tr. in Ajdukiewicz (1978), 140–153.

Ajdukiewicz, K. (1948). Epistemologia i semiotyka (Epistemology and semiotics). *Przegląd Filozoficzny*, 44, 336–347; Eng. tr. in Ajdukiewicz (1978), 182–191.

Ajdukiewicz, K. (1949). *Zagadnienia i kierunki filozofii.* Warszawa: Czytelnik; Eng. tr. (by H. Skolimowski, A. Quinton) *Problems & Theories of Philosophy*. Cambridge: Cambridge University Press.

Ajdukiewicz, K. (1978). *The Scientific World-Perspective and Other Essays, 1930–1963*. Dordrecht: Reidel.

Alston, W. (1996). *A Realist Conception of Truth*. Ithaca: Cornell University Press.

Armour, L. (1969). *The Concept of Truth*. Assen: Van Gorcum.

Armour-Garb, B. P., Beall, J. C. (Eds.) (2005). *Deflationary Truth*. Chicago: Open Court.

Armstrong, D. (2004). *Truth and Truthmakers*. Cambridge: Cambridge University Press.

Asser, G. (1972). *Einführung in die mathematische Logik*. Leipzig: Teubner.

Barnard, R., Ulatowski, J. (2016). Tarski's 1944 polemical remarks and naess' "experimental philosophy". *Erkenntnis*, 80, 457–477.

Båve, A. (2006). *Deflationism. A Use-Theoretic Analysis of the Truth-Predicates*. Stockholm: Stochkolm University.

Benmakhlouf, A. (Ed.) (2004). *Sémantique et épistemologie*. Casablanca: Editions Le Fennec.

Benacerraf, P. (1973). Mathematical truth. *The Journal of Philosophy*, 70, 661–679.

Benacerraf, P. (1996). What mathematical Truth – Could not Be – I. In Morton, Stich (1996), 9–59.

Benacerraf, P. (1998). What mathematical Truth – Could not Be – II. In Cooper, Truss (1998), 27–51.

Berger, U., Schichtenberg, H. (Eds.) (1999). *Computational Logic*. Berlin: Springer.

Beth, E. W. (1968). *The Foundations of Mathematics*. Amsterdam: North-Holland.

Black, M. (1948). The semantic definition of truth. *Analysis*, 8, 49–63; repr. (with additions), in Black (1949), 89–107, 251.

Black, M. (1949). *Language and Philosophy*. Ithaca: Cornell University Press.

Bueono, O., French, S. (2018). *Applying Mathematics. Immersion, Inference, Interpretation*. Oxford: Oxford University Press.

Burgess, A. G., Burgess, J. P. (2011). *Truth*. Princeton: Princeton University Press.

Butler, M. K. (2017). *Deflationism and Semantic Theories of Truth*. Manchester: Pendlebury Press.

Button, T., Walsh, S. (2018). *Philosophy and Model Theory*. Oxford: Oxford University Press.

Cat, J., Tuboly, A. T. (Eds.) (forthcoming). *New Perspectives on Otto Neurath's Life and Works*. Dordrecht: Springer.

Carnap, R. (1942). *Introduction to Semantics*. Cambridge, Mass.: Harvard University Press; 2nd ed., 1958.

Carnap, R. (1947). *Meaning and Necessity. A Study in Semantics of Modal Logic*. Chicago: University of Chicago Press; 2nd ed., 1956.

Carnap, R. (1950). *Logical Foundations of Probability*. Chicago: Chicago University Press; 2nd ed., 1962.

Cauman, L. S., Levi, I., Parsons, Ch., Schwartz, R. (Eds.) (1983). *How Many Questions? Essays in Honour of Sidney Morgenbesser*. Indianapolis: Hackett Publishing Company

Cieśliński, C. (2017). *The Epistemic Lightness of Truth. Deflationism and its Logic*. Cambridge: Cambridge University Press.

Cooper, B. S., Truss, J. K. (Eds.) (1998). *Sets and Proofs*. Cambridge: Cambridge University Press.

Couturat, L. (1905). *L'Algebre de la loqique*. Paris: Gauthier-Villars; Eng. tr., La Salle: Open Court 1914.

Da Costa, N. C. A., French, S. (2003). *Science and Partial Truth. A Unitary Approach to Models and Scientific Reasoning*. Oxford: Oxford University Press.

Dales, H. G., Oliveri, G. (Eds.) (1998). *Truth in Mathematics*. Oxford: Clarendon Press.

Dancy, J. (1985). *Introduction to Contemporary Epistemology*. Oxford: Basil Blackwell.

David, M. (1994). *Correspondence and Disquotation. An Essay on the Nature of Truth*. Oxford: Oxford University Press.

David, M. (2008). Tarski's Convention T and the concept of truth. In Patterson (2008), 133–156.

Davidson, D. (1967). Truth and meaning. *Synthese*, 17, 304–323; repr. in Davidson (1984), 37–54.

Davidson, D. (1970). Semantics for natural language. In *Linguaggio nella Socièta e nella Tecnica*. Milano: Edizioni di Communità, 12–22; repr. in Davidson (1984), 55–64.

Davidson, D. (1973). In defence of Convention T. In Leblanc (1973), 76–86; repr. in Davidson (1984), 65–75.

Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.

Dawson-Galle, P. (1998). *The Possibility of Relative Truth*. Aldershot: Ashgate.

Devitt, M. (1996). *Truth and Realism*. Princeton: Princeton University Press.

Dummett, M. (1978). *Truth and Other Enigmas*. Cambridge, Mass.: Harvard University Press.

Earmann, J. (1986). *A Primer of Determinism*. Dordrecht: Kluwer.

Etchemendy, J. (1990). *The Concept of Logical Consequence*. Cambridge, Mass.: Harvard University Press.

Field, H. (1972). Tarski's theory of truth. *Journal of Philosophy*, 69, 347–375; repr. with Postscipt in Field (2001), 3–29.

Field, H. (2001). *Truth and the Absence of Fact*. Oxford: Clarendon Press.

Frápolli, M. J. (2013). *The Nature of Truth: An Updated Approach to the Meaning of Truth Ascriptions*. Berlin: Springer.

García-Carpintero, M., Kölbel, M. (Eds.) (2008). *Relative Truth*. Oxford: Oxford University Press.

Gardiner, M. Q. (2000). *Semantic Challenges to Realism. Dummett and Putnam*. Toronto: University of Toronto Press.

Gårding, L. (1977). *Encounter with Mathematics*. Berlin: Springer.

Garrido, A., Wybraniec–Skardowska, U. (Eds.) (2018). *The Lvov–Warsaw School. Past and Present*. Basel: Birkhäuser.

Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, 23, 121–123.

Girard J.-Y. (1999). On the meaning of logical rules I: Syntax versus semantics. In Berger, Schichtenberg (1999), 215–272.

Girard, J.-Y. (2011). *The Blind Spot. Lectures on Logic*. Zürich.: European Mathematical Society.

Glanzberg, M. (2003). Minimalizm and paradoxes. *Synthese*, 135(1), 12–36.

Goldstern, M., Judah, H. (1995). *The Incompleteness Phenomenon. A New Course in Mathematical Logic*. Wellesley: A K Peters.

Greenough, P., Lynch, M. P. (Eds.) (2006). *Truth and Realism*. Oxford: Clarendon Press.

Haack, S. (1978). *Philosophy of Logics*. Cambridge: Cambridge Univeristy Press.

Halbach, V. (2011). *Axiomatic Theories of Truth*. Cambridge: Cambridge University Press.

Hilbert D., Bernays, P. (1939). Grundlagen der Mathematik, v. II. Berlin: Springer; 2nd ed., 1970.

Hill, Ch. (2002). *Thought and World. An Austere Portrayal of Truth, Reference, and Semantic Correspondence*. Cambridge: Cambridge University Press.

Hintikka, J., Czarnecki, T., Kijania–Placek, K., Placek, T., Rojszczak, A. (Eds.) (2003). *Philosophy and Logic. In Search of Polish Tradition. Essays in Honour of Jan Woleński on the Occasion of his 60th Birthday*. Dordrecht: Kluwer.

Hinzen, W. (1998). *The Semantic Foundations of Anti-Realism*. Berlin: Logos.

Horwich, P. (1998). *Truth*. Oxford: Clarendon Press; 2nd ed., 2000.

Horwich, P. (2005). *From a Deflationary Point of View*. Oxford: Clarendon Press.

Horwich, P. (2010). *Truth, Meaning, Reality*. Oxford: Clarendon Press.

Hugly, Ph., Sayward, Ch. (1996). *Intensionality and Truth. An Essay on the Philosophy of A. N. Prior*. Dordrecht: Kluwer.

Ingarden, R. (1925). Essentiale Fragen. Ein Beitrag zum Problem des Wesen. *Jahrbuch für Philosophie und philosopische Forschung*, 7, 125–304.

Ingarden, R. (1949). O sądzie warunkowym (On conditional proposition). *Kwartalnik Filozoficzny*, 18, 263–308; repr. in Ingarden (1972), 271–325.

Ingarden, R. (1972). *Z teorii języka i filozoficznych podstaw logiki* (Theory of Language andPhilosophical Foundations of Logic). Warszawa: Państwowe Wydawnictwo Naukowe.

Lewis, D. (Ed.) (1968). *British Philosophy, III*. London: George Allen and Unwin

Ketland, J. (1999). Deflationism and Tarski's paradise. *Mind*, 108, 69–94.

Ketland, J. (2005). Deflationism and the Gödel phenomena; reply to Tennant. *Mind*, 114, 75–88.

Khatchadourian, H. (1961). *The Coherence Theory of Truth. A Critical Evaluation*. Beirut: Americal University.

Kirkham, R. (1993). Tarski's physicalism. *Erkenntnis*, 38, 289–302.

Kokoszyńska, M. (1936). Über den Absoluten Wahrheitsbegriff und einige andere semantische Begriffe. *Erkenntnis*, 6, 143–165; repr. in Pearce, Woleński (1988), 276–292.

Kokoszyńska, M. (1948). What means a "Relativity" of "Truth". *Studia Philosophica*, III, 167–175.

Kokoszyńska, M. (1951). A refutation of the relativism of truth. *Studia Philosophica*, III, 93–149.

Kölbel, M. (2002). *Truth without Objectivity*. London: Routledge.

Kotarbiński, T. (1934). W sprawie pojęcia prawdy [Recenzja: A. Tarski, Pojecie prawdy w językach nauk dedukcyjnych] (On the Concept of Truth [Review of A. Tarski, The concept of truth in the languages of deductive sciences]). *Przegląd Filozoficzny*, 37, 85–91; repr. in Kotarbiński (1993), 252–258.

Kotarbiński, T. (1993). *Ontologia, teoria poznania i metodologia* (Ontology, Theory of Knowledge and Methodology). Wrocław: Zakład Narodowy im. Ossolińskich.

Krajewski, S. (1994). Tarski's Definition of Truth and Mathematics. In Twardowski, Woleński (1994), 16–33.

Künne, W. (2002). Disquotationalist conceptions of truth. *Schantz*, 2002, 176–193.

Künne, W. (2003). *Conceptions of Truth*. Oxford: Clarendon Press.

Kuźniar, A., Odrowąż–Sypniewska, J. (Eds.) (2016). *Uncovering Facts and Values. Studies in Contemporary Epistemology and Political Philosophy*. Leiden: Brill/Rodoppi.

Lakatos, I., Musgrave, A. (Eds.) (1968). *Proceedings of the International Colloquium in the Philosophy of Science, London, 1965*. Amsterdam: North-Holland.

Leblanc, H. (Ed.) (1973). *Truth, Syntax and Modality*. Amsterdam: North-Holland
Lowe, E. J., Rami, A. (Eds.) (2009). *Truth and Truth-Making*. Stocksfields: Acumen.
Magnani, L., Bertolloti, T. (Eds.) (2017). *Springer Handbook of Model-Based Science*. Berlin: Springer.
McGee, V. (1991). *Truth, Vagueness and Paradox*. Indianapolis: Hackett Publishing Company.
McGinn, C. (1980). Truth and use. In Platts (1980), 19–40.
McGrath, M. (2000). *Between Deflationism and Correpondence Theory*. New York: Garland.
Moreno, L. F. (1992). *Wahrheit und Korrespondenz bei Tarski. Eine Untersuchung der Wahrheitstheorie Tarskis als Korrepondenztheorie der Wahrheit*. Würzburg: Köningshausen & Neumann.
Morton, A., Stich, S. P. (Eds.) (1996). *Benacerraf and his Critics*. Oxford: Blackwell.
Mulligan, K., Simons, P. M., Smith, B. (1984). Truth-makers. *Philosophy and Phenomenological Research*, 44, 287–321; repr. in Lowe, Rami (2009), 59–86.
Naess, A. (1938). *"Truth" as Conceived by Those Who are not Professional Philosophers*. Oslo: I Kommisjon Hos Jacob Dybward.
Naess, A. (1953). *An Empirical Study of the Expression "True", "Perfectly Certain" and "Extremely Probable"*. Oslo: I Kommisjon Hos Jacob Dybward.
Neale, S. (2001). *Facing Facts*. Oxford: Clarendon Press.
Newman, A. (2002). *The Correspondence Theory of Truth. An Essay on the Metaphysics of Predication*. Cambridge: Cambridge: University Press.
Niiniluoto, I. (1987). *Truthlikeness*. Dordrecht: Kluwer.
Niiniluoto, I. (1994). Defending Tarski against his critics. In Twardowski, Woleński (1994), 48–68.
Niiniluoto, I. (1999). Theories of truth: Vienna, Berlin and Warsaw. In Woleński, Köhler (1999), 17–26.
Niiniluoto, I., Sintonen, M., Woleński, J. (Eds.) (2004). *Handbook of Epistemology*. Dodrecht: Kluwer.
Novak, Z., Simonyi, A. (Eds.) (2011). *Truth, Reference and Realism*. Budapest: CEU Press.
O,'Connor, D. J. O. (1975). *The Correspondence Theory of Truth*. London: Hutchinson.
O'Grady, P. (2002). *Relativism*. Montreal: McGill-Queen's University Press.
Olsson, E. J. (2005). *Against Coherence. Truth, Probability and Justification*. Oxford: Clarendon Press.
Pap, A. (1954). Propositions, sentences and the semantic concept of truth. *Theoria*, XX, 23–35.
Pataut, F. (Ed.) (2016). *Truth, Objects, Infinity. New Perspectives on the Philosophy of Paul Benacerraf*. Dordrecht: Springer.
Patterson, D. E. (2006). Tarski and the necessity reading of the convention T. *Synthese*, 151(1), 1–32.
Pearce, D., Woleński, J. (Eds.) (1988). *Logische Rationalismus. Philosophische Schriften der Lemberg–Warschauer Schule*. Frankfurt a. M.: Athenäum.
Pedersen, N. J., Wright, C. D. (2013). *Truth and Pluralism*. Oxford: Oxford University Press.
Platts, M. (Ed.) (1980). *Reference, Truth and Reality*. London: Routledge and Kegan Paul.
Popper, K. (1972). *Objective Knowledge. An Evolutionary Approach*. Oxford: Clarendon Press; repr. with additions, 1974.
Prior, A. N. (1971). *Objects of Thought*. Oxford: Clarendon Press.
Przełęcki, M. (1969). *The Logic of Empirical Theories*. London: Routledge and Kegan Paul.
Przełęcki, M. (1974). Some philosophical consequences of the semantic definition of truth. *Dialectic and Humanism*, I, 117–128.
Przełęcki, M. (1974a). On model-theoretic approach to interpretation of empirical theories. *Synthese*, 26, 401–406.
Putnam, H. (1983). On truth. In Cauman, Levi, Parsons, Schwartz (1983), 35–56; repr. in Putnam (1994), 315–329.
Putnam, H. (1985–1986). On comparison of something with something else. *New Literary History*, 17, 61–79; repr. in Putnam (1994), 330–350.
Putnam, H. (1994). *World and Life*. Cambridge, Mass.: Harvard University Press.

Quine, V. V. O. (1953). *From a Logical Point of View*. Cambridge, Mass.: Harvard University Press.

Quine, V. V. O. (1953a). Notes on the theory of reference. In Quine (1953), 130–138.

Rami, A. (2009). *Wahrheit und Deflation. Eine kritische Untersuchung deflationärer Wahrheitskonzeptionen*. Paderborn: Mentis.

Rami, A. (2009a). Introduction: truth and truth-making. Loewe, Rami, 2009, 1–36.

Rasmussen, J. (2014). *Defending the Correspondence Theory of Truth*. Cambridge: Cambridge University Press.

Reboul, A. (Ed.) (2014). Mind, Values, and Metaphysics. Philosophical Essays in Honor of Kevin Mulligan, v. 1. Dordrecht: Springer.

Rescher, N. (1973). *The Coherence Theory of Truth*. Oxford: Clarendon Press.

Rescher, N. (1985). Truth as ideal coherence. *The Review of Metaphysics*, 38, 795–806.

Russell, B. (1948). *Human Knowledge. Its Scope and Limits*. London: George Allen and Unwin.

Russell, B. (1984). *The Theory of Knowledge. The 1913 Manuscript*. London: George Allen and Unwin.

Scaglia, B. (2011). *The Epistemology of Truth: Veritas, Logic and Interpretation*. Memphis: Books LLC.

Schantz, R. (Ed.) (2002). *What is Truth?* Berlin: de Gruyter.

Schlick, M. (1931). Das Kausalität in den gegenwärtigen Physik. *Naturwisseschaften*, 19, 145–162; Eng. tr. in Schlick (1979), 176–209.

Schlick, M. (1979). Collected Papers, v. II. Dordrecht: Reidel.

Schröter, J. (1996). *Zur Meta-theorie der Physik*. Berlin: de Gruyter.

Shope, R. K. (1983). *The Analysis of Knowledge. A Decade of Research*. Princeton: Princeton University Press.

Sinnott-Armstrong, W., Raffman, D., Asher, N. (Eds.) (1995). *Modality, Morality and Belief. Essays in Honor of Ruth Barcan Marcus*. Cambridge: Cambridge University Press.

Simons P. (2003). Absolute truth in a changing world. In Hintikka, Czarnecki, Kijania–Placek, Placek, Rojszczak (2003), 37–54.

Soames, S. (1995). T-Sentences. In Sinnott-Armstrong, Raffman, Asher (1995), 250–270.

Soames, S. (1999). *Understanding Truth*. Oxford: Oxford University Press.

Stegmüller, W. (1957). *Das Wahrheitsbegriff und die Idee der Semantik*. Wien: Springer.

Suppes, P. (2002). *Representation and Invariance of Scientific Structures*. Stanford: CSLI Publications.

Suszko, R. (1968). Formal logic and the development of knowledge. In Lakatos (1968), 210–222.

Tarski, A. (1933). *Pojęcie prawdy w językach nauk dedukcyjnych* (The Concept of Truth in Languages of Deductive sciences). Warszawa: Towarzystwo Naukowe Warszawskie; Germ. tr. (with additions) In Tarski (1935).

Tarski, A. (1935). *Der Wahrheitsbegriff in den formalisierten Sprachen. Studia Philosophica*, 1, 261–405; repr. in Tarski (1986), v. 2, 51–198; Eng. tr. *The Concept of Truth in Formalized Languages*, in Tarski (1956), 152–278 (page-references Tarski (1933)).

Tarski, A. (1936). Grundlegung der wissenschaftliche Semantik. In *Actes du Congès International de Philosophie Scientifique*, v. 3, 1–8. Paris: Herman; Eng. tr. in Tarski (1956), 401–408.

Tarski, A. (1944). Semantic theory of truth and the foundations of semantics. *Philosophy and Phenomenological Research*, 4, 341–395; repr. in Tarski (1986), v. 2, 661–697.

Tarski, A. (1956). *Logic, Semantics, Metamathematics. Papers of 1923 to 1938* (tr. by J. H. Woodger). Oxford: Clarendon Press; 2nd ed., Indianapolis: Hackett Publishing Company 1983.

Tarski, A. (1969). Truth and proof. *Scientific American*, 220(6), 63–77; repr. in Tarski (1986), v. 4, 399–423.

Tarski, A. (1986). Collected Papers, 1–4, Basel: Birkhäuser.

Taylor, B. (2006). *Truth and Realism*. Oxford: Oxford University Press.

Tennant, N. (1987). *Anti-Realism and Logic. Truth as Eternal*. Oxford: Clarandon Press.

Tennant, N. (1997). *The Taming of the True*. Oxford: Clarendon Press.

Tennant, N. (2002). Deflationism and the Gödel Phenomena. *Mind*, 111, 551–582.

Tennant, N. (2005). Deflationism and the Gödel Phenomena: Reply to Ketland. *Mind*, 114, 89–96.
Twardowski, B., Woleński, J. (Eds.) (1994). *60 Years of Tarski's Definition of Truth*. Kraków: Phileb.
Udemadu, K. (1995). *Einwände gegen Alfrer tarski semantische Theorie der Wahrheit*. Norderstedt: Grin.
Ulatowski, J. (2016). Ordinary truth in Tarski and Næss. In Kuźniar, Odrowąż–Sypniewska (2016), 67–89.
Ulatowski, J. (2017). *Commonsense Pluralism about Truth. An Empirical Defence*. Basinstoke: Palgrave Macmillan.
Vision, G. (2004). *Veritas: The Correspondence Theory and Its Critics*. Cambridge, Mass.: The MIT Press.
Vuillemin, J. (1996). *Necessity or Contingency. The Master Argument*. Stanford: CSLI Publications.
Waismann, F. (1956). How I see philosophy? In Lewis (1956), 447–490; repr. in Waismann (1968), 1–38.
Waismann, F. (1965). *The Principles of Linguistic Philosophy*. London: Macmillan.
Waismann, F. (1968). *How I See Philosophy*. London: Macmillan.
Waismann, F. (1976). *Logik, Sprache, Philosophie*. Stuttgart: Reclam.
Walker, R. C. (1989). *The Coherence Theory of Truth. Realism, Anti-Realism, Idealism*. London: Routledge.
Weingartner, P. (2000). *Basic Questions on Truth*. Dordrecht: Kluwer.
Wessel, H. (1999). *Logik und Philosophie*. Berlin: Logos.
Williams, C. J. F. (1976). *What is Truth?*. Cambridge: Cambridge University Press.
Williamson, T. (2013). *Modal logic as Metaphysics*. Oxford: Oxford University Press.
Wójcicki, R. (1979). *Topics in the Formal Methodology of Empirical Science*. Dordrecht: Reidel.
Woleński, J. (1989a). Brentano's criticism of the correspondence theory of truth and Tarski's semantic theory. *Topoi*, 6, 105–110; repr. in Woleński (1999), 103–109.
Woleński, J. (1999). *Essays in the History of Logic and Logical Philosophy*. Kraków: Jagiellonian University Press.
Woleński, J. (1999b). The semantic revolution—Rudolf Carnap, Kurt Gödel, Alfred Tarski. In Woleński, Köhler (1998), 1–15.
Woleński, J. (2001). In defense of the semantic definition of truth. *Synthese*, 126(1–2), 67–90.
Woleński, J. (2004). Logic, semantics and realism. In Benmakhlouf (2004), 135–148; repr. in Woleński (2011), 51–60.
Woleński, J. (2004g). Analytic vs. Synthetic and A Priori vs. A Posteriori. In Niiniluoto, Sintonen, Woleński (2004), 781–839.
Woleński, J. (2007a). Two critical contributions to the problem of truth and meaning. *Polish Journal of Philosophy*, 1(2007), 137–141; repr. in Woleński (2011), 183–187.
Woleński, J. (2011). *Essays on Logic and Its Applications in Philosophy*. Frankfurt a. M.: Peter Lang.
Woleński, J. (2012). Do we need to reform the old T-scheme. *Discusiones Filozóficas*, 20(13), 73–85; repr. in Woleński (2018), 45–50.
Woleński, J. (2014). Truth-makers and convention T. In Reboul (2014), 79–84; repr. in Woleński (2018), 155–160.
Woleński, J. (2017). Semantic definition of truth, empirical theories and scientific realism. In Agazzi (2017), 279–294.
Woleński, J. (2018). *Logic and Its Philosophy*. Frankfurt a. M.: Peter Lang.
Woleński, J. (2018b). Some philosophical aspects of semantic theory of truth. In Garrido, Wybraniec–Skardowska (2018), 371–389.
Woleński, J. (2018). *Essays on Logic and Its Philosophy*. Frankfurt a. M.: Peter Lang.
Woleński, J., Köhler, E. (Eds.) (1999). *Alfred Tarski and the Vienna circle. Austro-Polish Connection in Logical Empiricism*. Dordrecht: Kluwer.

Conclusion

If we look at **SDT** within metalogic, it functions as a normal definition in the sense accepted in the traditional or contemporary theory of definition. Its preparatory part concerns the concept of satisfaction and is inductive—the ultimate definition has the form of equivalence. On the other hand, the status of **SDT** inside philosophy is fairly problematic and calls for further explanations. In fact, definitions in philosophy do not have a good reputation, and this situation has very convincing reasons. So-called analytic definitions have to cover so many intuitions, frequently—mutually conflicting or non-homogenic, that their explanatory effectiveness is very limited, particularly, as successful means of exhibiting meanings of relevant concepts (think on the notion of realism, for example, but many other suggestive illustrations immediately comes to mind). On the other hand, if someone proposes a synthetic or regulative definition, a natural question is why some intuition is chosen as fundamental, whereas others are neglected or considered as secondary, or why a normalized decision is made at the given point but not at a different place; moreover, regulative definitions are always conventional to some extent and subjected to the objection that they ignore the essence (a beloved category of philosophers) of defined concepts or entities. Moreover, a chance for producing a short (of course, taking this property *cum grano salis*) formulation falling under the scheme '*X* is *Y*' where the letter *X* refers to a definiendum, but the letter *Y*—the definiens associated with what is defined, seems rather small. This conclusion is sufficiently confirmed by the entire history of philosophy.

Taking into account these circumstances as sufficiently indicating various difficulties in using definitions in doing philosophy (I do not suggest that philosophy is exceptional in this respect, because the similar situation occurs in the humanities as well as it is documented by various attempts to define art, religion, music, language, etc.). I guess that other analytic procedure has a greater significance for philosophical enterprise. I think here about Carnap's idea of explication (see Carnap 1950, pp. 3–8) as a much better servant for philosophical needs than formulas having the standard shape of definitions in the sense of elementary (school) logic. The method of explication consists in using and comparing two ingredients, namely

J. Woleński, *Semantics and Truth*, Logic, Epistemology, and the Unity of Science 45, https://doi.org/10.1007/978-3-030-24536-8

an explicandum (what is explicated in a given conceptual situation; analogy to definiendum) and an explicans (by which the explicatum is explicated—analogy to definiens; one can say that explicatum is clarified, explained, illuminated, specified, etc.). I would add at this place that interpretative consequences (see Introduction) have much more to do with explications in Carnap's sense that with definitions. In fact, I consider my investigations undertaken in the book as interpretative consequences derived (via hermeutics which I tried to exhibit in an explicit manner) from a well-established logical theory.

Truth in the ordinary meaning, or even scientific is the explicandum on which the present book focuses (some authors speak of the pre-theoretical sense of being true in this context), but **STT** (not only **SDT**) functions as the actual explicatum. Carnap formulated the following constraints (conditions of adequacy) for the procedure of explicating (they pertain to properties of explicanda):

(a) similarity to the explicandum;
(b) exactness;
(c) fruitfulness;
(d) simplicity.

Although (a)–(d) are highly evaluative (it is nothing strange in philosophy), Carnap himself believed that, they might serve as tools for transforming philosophy into science (his analysis addressed the concept of inductive probability as explicated by the mathematical axiomatic theory of probability). I have no such ambitions vis-a-vis the concept of truth, and my task aims at a philosophical clarification. On the other hand, the points (a)–(d) can be also used in evaluation of this more moderate intention. Checking whether **STT** satisfies Carnap's constraints one can observe:

Ad (a) In (**DG4IX**), I argued that **STT** is close to the ordinary usage of 'is true', even more closer than other popular philosophical theories of truth. I consider this fact as a very serious advantage of the semantic theory of truth;

Ad (b) **STT** is exact owing to its formulation in logical and metalogical terms;

Ad (c) **STT** enables us to clarify many problems of traditional aletheiology;

Ad (d) **STT** is simple—at least for those who know contemporary logic.

When I studied philosophy with Roman Ingarden, he, as a faithful phenomenologist, denied that logic could have a deeper philosophical significance. According to him, **SDT** should be interpreted as a logical construction, relevant for logic, but not for philosophy. Is **STT** logical or not? Well, Tarski's definition does not belong to pure logic (in the understanding outlined in Chap. 5), but to metalogic. I included in previous chapters a considerable amount of logic and metalogic (or metamathematics) in order to show that even if the semantic theory of truth is not regarded as the best, it has sufficient virtues to be considered as a serious philosophical doctrine. However, I feel myself obliged to say that most statements in this book should be taken as a piece of philosophical speculation built on logic. Finally, I would like point out still one virtue of logical analysis. As I noted at the end of Introduction, Tarski is not sacrosanct. The same concerns **STT**. Although I defend this theory, I am fully aware that it is opened to criticism. The virtue in question consists in the fact that logic makes criticism easier than other philosophical devices, because the former forces transparency of reasoning and its results.

INTEGRATED BIBLIOGRAPHY

(Names of translators are omitted)

Abeles, F., Fuller, M. E. (Eds.) (2016). *Modern Logic 1850–1950. East and West.* Basel: Birkhäuser.

Achourioti, Th., Galinon, H., Fernández, J. M., Fujimoto, K. (Eds.) (2015). *Unifying the Philosophy of Truth*. Dordrecht: Springer.

Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E. Romanowska, A, Woleński, J. (Eds.) (2004). *Provinces of Logic Determined. Essays in the Memory of Alfred Tarski. Parts I, II and III* (*Annals of Pure and Applied Logic*, 127). Amsterdam: Elsevier.

Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E. Romanowska, A, Woleński, J. (Eds.) (2004a). *Provinces of Logic Determined. Essays in the Memory of Alfred Tarski. Parts VI, IV and VI* (*Annals of Pure and Applied Logic*, 127). Amsterdam: Elsevier.

Aersten, J. A. (1996). *Medieval Philosophy & the Transcendentals*. Leiden: Brill.

Agazzi, E. (Ed.) (1981). *Modern Logic*. Dordrecht: Reidel.

Agazzi, E. (Ed.) (2017). *Varietes of Scientific Realism*. Dordrecht: Springer.

Ajdukiewicz, K., (1931). O znaczeniu wyrażeń (On the meaning of expressions). In *Księga Pamiątkowa Polskiego Towarzystwa Filozoficznego we Lwowie*. Lwów: Polskie Towarzystwo Filozoficzne, 31–77; Eng. tr. in Ajdukiewicz (1978), 1–34.

Ajdukiewicz, K. (1935). Die syntaktische Konnexität. *Studia Philosophica*, 1, 1–27; Eng. tr. in Ajdukiewicz (1978), 118–139.

Ajdukiewicz, K. (1937). Problemat transcendentalnego idealizmu w sformułowaniu semantycznym (A semantical version of the problem of transcendental idealism). *Przegląd Filozoficzny*, **40**, 271–287; Eng. tr. in Ajdukiewicz (1978), 140–153.

Ajdukiewicz, K. (1948). Epistemologia i semiotyka (Epistemology and semiotics). *Przegląd Filozoficzny*, 44, 336–347; Eng. tr. in Ajdukiewicz (1978), 182–191.

Ajdukiewicz, K. (1949). *Zagadnienia i kierunki filozofii.* Warszawa: Czytelnik; Eng. tr., *Problems & Theories of Philosophy*. Cambridge: Cambridge University Press.

Ajdukiewicz, K. (1964). Zagadnienie empiryzmu a koncepcja znaczenia (Empiricism and the problem of meaning). *Studia Filozoficzne,* 1(36), 3–14; Eng.tr . in Ajdukiewicz 1978, 306–319.

Ajdukiewicz, K. (1978). *The Scientific World-Perspective and Other Essays, 1930–1963*. Dordrecht: Reidel.

Ajdukiewicz, K. (1993). Kategorie syntaktyczne i antynomie logiczne. Wykłady z semantyki logicznej w r. a. 1930/1931 (Syntactic categories and logical antinomies. Lectures on logical semantics in the academic year 1930/1931). *Filozofia Nauki* I/1, 163–182.

Albuquerque, D. (1995). *Truth and Action in Speech Acts. A Study on J. L. Austin's Essay on Truth*. New Delhi: Intercultural Publications.

Allan, K. (Ed.) (2009). *The Concise Encyclopedia of Semantics.* Amsterdam: Elsevier.

J. Woleński, *Semantics and Truth*, Logic, Epistemology,
and the Unity of Science 45, https://doi.org/10.1007/978-3-030-24536-8

Anderson, C. A., Zelëny, M. (Eds.) (2001). *Logic, Meaning and Computation*. Dordrecht: Kluwer.

Alston, W. (1996). *A Realist Conception of Truth.* Ithaca: Cornell University Press.

Andréka, H., Madárasz, J., Németi, I. (2003). Why first-order logic? (unpublished).

Armour, L. (1969). *The Concept of Truth.* Assen: Van Gorcum.

Armour-Garb, B. P., Beall, J. C (Eds.) (2005). *Deflationary Truth.* Chicago: Open Court. Anderson, C. A., Zeleny, M. (Eds.) (2001), *Logic, Meaning, and Computation: Essays in Memory of Alonzo Church.* Kluwer: Dordrecht.

Armstrong, D. (2004). *Truth and Truthmakers*. Cambridge University Press: Cambridge

Assmann, J. (1990). *Ma'aat. Gerechtigkeit und Unsterblichkeit in Alten Ägypten*. München: C. H. Beck.

Asser, G. (1972). *Einführung in die mathematische Logik*. Leipzig: Teubner.

Auxier, R. E., Hahn, L. E. (Eds.) (2006). *The Philosophy of Jaakko Hintikka*. La Salle: Open Court.

Ayer, A. J. (1946). *Language*, *Truth and Logic*, 2nd ed. London: Gollancz.

Ayer, A. J. (1977). *A Part of My Life*. London: Collins.

Baldwin, T. (1991). Can there be a substantive account of truth? In Cooper, Engel (1991), 22–39.

Bar-Am, N., Gattei, S. (Eds.) (2017). *Encouraging Opnness. Essays for Joseph Agassi on the Occasion of His 90th Birthday*. Dodrecht: Springer.

Bar-Hillel, Y. (1970). *Aspects of Language.* Amsterdam: North–Holland

Bar-Hillel, Y. (1970a). Husserl's conception of a purely logical grammar. In Bar-Hillel (1970), 89–97.

Barnard, R, Ulatowski, J. (2016). Tarski's 1944 polemical remarks and Naess' "experimental philosophy". *Erkenntnis*, 80, 457–477.

Barnes, J. (1993). Meaning, saying and thinking. In Döring, Ebert 1993, 47–61.

Barnes, J. (2007). *Truth, etc. Six Lectures on Ancient Logic.* Oxford: Clarendon Pres.

Bartlett, S. J. (Ed.) (1992). *Reflexivity. A Source-Book in Self-Reference.* Amsterdam: North-Holland.

Bartlett, S. J., Suber, P. (Eds.) (1987). *Self-Reference. Reflections on Reflexivity.* The Hague: Nijhoff.

Barwise, J. (1975). *Admissible Sets and Structures*. Berlin: Springer.

Barwise, J. (1985). Model-theoretic logics: background and aims. In Barwise, Feferman (1985), 3–23.

Barwise J., Etchemendy J. (1987). *The Liar: An Essay on Truth and Circularity*. Oxford: Oxford University Press.

Barwise, J., Feferman, S. (Eds.) (1985). *Model-Theoretic Logics*. Berlin: Springer.

Barwise, K., Perry, J. (1983). *Situations and Attitudes.* Cambridge, Mass.: The MIT Press.

Båve, A. (2006). *Deflationism. A Use-Theoretic Analysis of the Truth-Predicates*. Stockholm: Stochkolm University.

Bärthlein, K. (1972). *Der Transzendentalehre der alten Ontologie*, Teil I: *Der Transzendentallehre im Corpus Aristotelicum*. Berlin: de Gruyter.

Beall, J. C., Armour-Garb (Ed.) (2009). *Deflationism and Paradox.* Oxford: Clarendon Press.

Beall, J. C, Restall, G. (2006). *Logical Pluralism.* Clarendon Press: Oxford.

Beal, J. C., Glanzberg, M. Ripley, D. (2018). *Formal Theories of Truth*. Oxford: Oxford University Press.

Beeh, V. (2003). *Die halbe Wahrheit. Tarskis Definition& Tarskis Theorem*. Paderborn: Mentis.

Bello, A. G. A. (2004). Some methodological controversies in African philosophy. In Wiredu (2004), 263–273.

Bell, J., Machover, M. (1977). *A Course in Mathematical Logic.* Amsterdam: North-Holland.

Benacerraf, P. (1973). Mathematical truth. *The Journal of Philosophy*, 70, 661–679.

Benacerraf, P. (1996). What mathematical truth – could not be – I. In Morton, Stich 1996, 9–59.

Benacerraf, P. (1998). What mathematical truth –could not be – II. In Cooper, Truss 1998, 27–51.

Benmakhlouf, A. (Ed.) (2004). *Sémantique et épistemologie*. Casablanca: Editions Le Fennec.

Bennet, B. M., Hoffman, D. D., Prakash Ch. (1989). *Observer Mechanics. A Formal Theory of Perception.* New York: Academic Press.

Berger, U, Schichtenberg, H. (Ed.) (1999). *Computational Logic*. Berlin: Springer.

Beth, E. W. (1962). *Formal Methods. An Introduction to Symbolic Logic and to the Study of Effective Operations*. Dordrecht: Reidel.

Beth, E.W. (1968). *The Foundations of Mathematics.* Amsterdam: North-Holland.

Betti, A. (2004). Leśniewski's early liar, Tarski and natural language. In Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E. Romanowska, A, Woleński, J. (2004a), 267–287.

Beziau, J.-Y. (Ed.) (2007). *Logica Universalis. Toward a General Theory of Logic*. Basel: Birkhäuser.

Beziau, J.-Y., Carnielli, W., Gabbay, D. (2007). *Handbook of Paraconsistency.* London: College Publications.

Black, M. (1948). The semantic definition of truth. *Analysis*, 8, 49–63; repr. (with additions), in Black (1949), 89–107, 251.

Black, M. (1949). *Language and Philosophy*. Ithaca: Cornell University Press.

Boeder, H. (1959). Der frügrechische Wortgebrauch von Logos und Aletheia". *Archiv für Begriffsgeschichte*, 4, 82–112.

Boehner, Ph. (1958). *Collected Articles on Ockham*. New York: The Franciscan Institute St. Bonaventure.

Bolander, Th., Hendricks, V. F., Pedersen, S. A. (Eds.) (2006). *Self-Reference.* Stanford: CSLI Publications.

Bolzano, B. (1837). *Wissenschaftslehre.* Sulzbach: Geiblish; Eng. tr., *Theory of Science.* Oxford: Oxford University Press 2014.

Boolos, G. (1975). On second-order logic. *Journal of Philosophy*, 72, 509–527; repr. in Shapiro (1996), 70–87.

Borkowski, L. (1991). Characterization of quantifiers in the axiomatic theory of consequence. In Borkowski, Stępień (1991), 37–39.

Borkowski, L., Stępień, A. B. (Eds.) (1991). *Studies in Logic and the Theory of Knowledge 2*. Lublin: Towarzystwo Naukowe KUL.

Bosanquet, B. (1911). *Logic or the Morphology of Knowledge*. Oxford: Clarendon Press.

Bradley, F. H. (1883). *The Principles of Logic*. Oxford: Clarendon Press.

Bradley F. H. (1893). *Appearance and Reality*. *A Metaphysical Essays*. Oxford: Clarendon Press.

Bradley, F. H. (1907). On truth and copying. *Mind*, XVI, 94–110; repr. in Bradley (1914), 107–116.

Bradley, F. H. (1914). *Essays on Truth and Reality*.Oxford: Clarendon Press.

Bradley, F. H. (1922). *The Principles of Logic*, 2nd ed. Oxford: Clarendon Press.

Brandl, J. (2017). Was Brentano an early Deflationist about Truth? *The Monist*, 100(1), 1–14.

Brentano, F. (1930), *Wahrheit und Evidenz.* Leipzig: Felix Meiner; Eng. tr., *The True and the Evident.* London: Routledge and Kegan Paul 1966.

Bréal, M. (1997). Essai de sémantique: science des significations. Hachette: Paris.

Bueono, O., French, S. (2018). *Applying Mathematics. Immersion, Inference, Interpretation.* Oxford University Press: Oxford.

Burks, A. (1966). Editior's comments. In Von Neumann (1966), 53–56.

Burnet, J. (1892). *Early Greek Philosophy*. London: Adam and Charles Black.

Brożek, A., Chybińska, A., Jadacki, J. Woleński, J. (Eds.) (2015). *Tradition of the Lvov–Warsaw School. Ideas and Continuations*. Leiden: Brill/Rodopi.

Buldt, B. (2002). Philosophische Implikationen der Gödelsche Sätze? Ein Bericht. In Köhler et al. (2002), v. 2, 395–438.

Bultmann, R. (1933). Der griechische und hellenistische Sprachgebrauch von ἀλήθεια". In Kittel 1933, 239–240.

Burgess, A. G., Burgess, J. P. (2011). *Truth.* Princeton: Princeton University Press.

Butler, M. K. (2017). *Deflationism and Semantic Theories of Truth*. Manchester: Pendlebury Press.

Button, T., Walsh, S. (2018). *Philosophy and Model Theory.* Oxford: Oxford University Press.

Candlish, S. (2007). *The Russell/Bradley Dispute and its Significance for Twentieth Century Philosophy*.Basingstoke: Palgrave Macmillan.

Cann, R. (1993). *Formal Semantics. An Introduction*. Cambridge: Cambridge University Press.

Cantone, D., Omodeo, E., Policriti, A. (2018). *Set Theory for Computing. From Decision Procedures to Declarative Programming with Sets*. Berlin: Springer.

Carnap, R. (1934). *Logische Syntax der Sprache.* Wien: Springer; Eng. tr., London: Routledge and Kegan Paul 1937.

Carnap, R. (1934a). Die Antinomien und die Unvollständingkeit der Mathematik. *Monatshefte für Mathematik und Physik*, 41, 263–284.

Carnap, R. (1939). *Foundations of Logic and Mathematics.* Chicago: University of Chicago Press.

Carnap, R. (1942). *Introduction to Semantics.* Cambridge, Mass: Harvard University Press; 2nd ed., 1958.

Carnap, R. (1947). *Meaning and Necessity. A Study in Semantics of Modal Logic*. Chicago: University of Chicago Press; 2nd ed. 1956.

Carnap, R. (1950). *Logical Foundations of Probability.* Chicago: Chicago University Press; 2nd ed., 1962.

Carnap, R. (1963). Intellectual autobiography. In Schilpp (1963), 3–84.

Carretero, F. M. L. (1983). *Der aristotelische Wahrheitsbegriff und die Aufgabe der Semantik.* Köln: Universität zu Köln.

Cartwright, R. (1987). A neglected theory of truth. In R. Cartwright, *Philosophical Essays*, Cambridge, Mass.: The MIT Press, 71–93.

Cat, J., Tuboly, A. T. (Eds.) (forthcoming). *New Perspectives on Otto Neurath's Life and Works.* Dordrecht: Springer.

Casari, E. (2006). *La matematica della verità. Strumenti matematici della semantica logica*. Torino: Bollati.

Casari, E. (2016). *Bolzano's Logical System*. Oxford: Oxford University Press.

Cauman, L. S., Levi, I., Parsons, Ch., Schwartz, R. (Eds.) (1983) *How Many Questions? Essays in Honour of Sidney Morgenbesser*. Indianapolis: Hackett Publishing Company.

Cavini, W. (1993). Chrysippus on speaking truly and the liar. In Döring , Ebert 1993, 85–109.

Chang, C. C., Keisler, H. J. (1979). *Model Theory.* Amsterdam: North-Holland.

Changeux, J.–P. (2002). *Physiology of Truth. Neuroscience and Human Knowledge.* Cambridge, Mass.: The Belknap Press.

Chappell, V. C. (Ed.) (1964). *Ordinary Language*. Englewood-Clifts: Prentice-Hall.

Church, A. (1943). Carnap's *Introduction to Semantics*. *The Philosophical Review*, 52, 298–305; repr. in Church (2019), 259–264. .

Church A. (1956). *Introduction to Mathematical Logic*. Princeton: Princeton University Press.

Church, A. (1976). Comparison of Russell's resolution of the semantical nntinomies with that of Tarski. *The Journal of Symbolic Logic*, 41, 747–760; repr. in Martin (1984), 291–306 and in Church (2019), 794–806.

Church, A. (2019). *Collected Papers.* The MIT Press.

Cieśliński, C. (2017). *The Epistemic Lightness of Truth. Deflationism and its Logic*. Cambridge: Cambridge University Press.

Classen, C. J. (1989). Protagoras' *aletheia*. In Huby, Neal (1989), 13–38.

Coffa, A. (1987). Carnap, Tarski and search for truth. *Nous*, 21, 547–572.

Coffa, A. (1991). *The Semantic Tradition from Kant to Carnap. To the Vienna Station*. Cambridge: Cambridge University Press.

Conford, F. M. (1935). *Plato's Theory of Knowledge. The Theaetetus and the Sophist of Plato with a Running Commentary*. London: Kegan Paul, Trench, Trubner.

Conford, F. M. (1952). *Principium Sapientiae: The Origin of Greek Philosophical Thought*. Cambridge: Cambridge University Press.

Cook, R. T. (2014). *The Yablo Paradox. An Essay on Circularity.* Oxford: Oxford University Press.

Cooper, N., Engel, P. (Eds.) (1991). *New Inquiries into Meaning and Truth.* New York: St. Martin's Press.

Cooper, B. S., Truss, J. K. (Eds.) (1998). *Sets and Proofs*. Cambridge: Cambridge University Press.

Copi, I. (1971). *The Theory of Logical Types.* London: Routledge and Kegan Paul.

Corcoran, J. 2001. Second-order logic. In Anderson, Zelëny, 61–75.

Couturat, L. (1905). *L'Algebre de la loqique.* Paris: Gauthier-Villars; Eng. tr., La Salle: Open Court 1914.

Crivelli, P. (2004). *Aristotle on Truth.* Cambridge: Cambridge University Press.

Czarnocka, M. (2017). *A Path to a Conception of Symbolic Truth.* Frankfurt am Main: Peter Lang.

Czeżowski, T. (1919). Imiona i zdania (Names and sentences). *Przegląd Filozoficzny*, 22, 1–19.

Da Costa, N. C. A., French, S. (2003). *Science and Partial Truth. A Unitary Appproach to Models and Scientific Reasoning*. Oxford: Oxford University Press.

Dales, H. G., Oliveri, G. (Eds.) (1998). *Truth in Mathematics.* Clarendon Press: Oxford.

Dancy, J. (1985). *Introduction to Contemporary Epistemology.* Oxford: Basil Blackwell.

David, M. (1994). *Correspondence and Disquotation. An Essay on the Nature of Truth.* Oxford: Oxford University Press.

David, M. (2004). Theories of truth. In Niiniluoto, Sintonen, Woleński (2004), 331–414.

David, M. (2008). Tarski's Convention T and the concept of truth. In Patterson (2008), 133–156.

Davidson, D. (1967). Truth and meaning. *Synthese*, 17, 304–323; repr. in Davidson (1984), 37–54.

Davidson, D. (1970). Semantics for natural language. In *Linguaggio nella Socièta e nella Tecnica*. Milano: Edizioni di Communità,12–22; repr. in Davidson (1984), 55–64

Davidson, D. (1969). True to facts. *The Journal of Philosophy*, 66, 748–764; repr. in Davidson 1984, 37–54.

Davidson, D. (1973). In defence of Convention T. In Leblanc (1973), 76–86; repr. in Davidson (1984), 65–75.

Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.

Davidson, D. (1996). The folly to trying to define truth. *Dialogue and Universalism*, 6, 39–53.

Davis, M. (Ed.) (1965). *The Undecidable. Basic Papars on Undecidable Propositions, Unsolvable Problems And Computable Functions.* Hewlett, N. Y.: Raven Press.

Dawson, J. W. (1996). *Logical Dilemmas: The Life and Work of Kurt Gödel*. Wellesley, Mass.: A. K. Peters.

Dawson-Galle, P. (1998). *The Possibility of Relative Truth.* Aldershot: Ashgate.

Detienne, M. (1967). *Les Maîtres de vérité dans la grèce archaïque*. Paris: François Mason. Eng. tr., *The Masters of Truth in Archaic Greece*. Cambridge, Mass.: The MIT Press 1996.

Devitt, M. (1984). *Realism and Truth.* Princeton: Princeton University Press; 2nd ed., 1991.

De Fioro, C. (2013). *La forma della verità. Logica e filosofia nell'opera di Alfred Tarski*. Milano: Mimesis.

De Rijk, L. M. (Ed.) (1956). *Petrus Abaelardus Dialectica*. Assen: Van Gorcum.

De Rijk, L. M. (1956a). Introduction. In De Rijk (1956), I–XCVII.

Dégremont, C. , Keiff, L., Rückert, H. (Eds.) (2008). *Dialogues, Logics and Other Strange Things. Essays in Honour of Shahid Rahman*. London: College Publications.

Doets, K. (1996). *Basic Model Theory*. Stanford: Center for the Study of Language and Information.

Dodd, J. (2000). *The Identity Theory of Truth.* London: Macmillan.

Döring, K., Ebert, Th. (Eds.) 1993. *Dialektiker und Stoiker. Zur Logik der Stoa und ihrer Vorläufer*. Stuttgart: Franz Steiner Verlag.

Dreben, B., Van Heijenoort, J. (1986). Introductory note. In Gödel (1986), 44–59.

Dumitriu, M., Sandu, G. (Eds.) (2013). *Truth*. Bucureşti: Editura Universitãtii din Bucureşti.

Dummett, M. (1978). *Truth and Other Enigmas*. Cambridge, Mass.: Harvard University Press.

Edwards, D. (2018). *The Metaphysics of Truth.* Oxford University Press: Oxford.

Edwards, P. (1967). *The Encyclopedia of Philosophy*, v. 1–8. New York: Macmillan.

Eggensperger, Th, Engel, U. (Eds.) (1995). *Wahrheit. Versuche zwischen Hochscholastik und Postmoderne*. Mainz: Matthias-Grünenwald Verlag.

Eisler, R. (1930). *Wörterbuch der philosophischen Begriffe*, v. III. Berlin: Mittler.

Enderton, H. B. (1972). *A Mathematical Introduction to Logic*. New York: Academic Press.

Enders, M. (1999). *Wahrheit und Notwendigkeit. Die Theorie der Wahrheit bei Anselm of Canterbury im Gesamtzusammenhang seines Denkens und unter besonderer Berücksichtigung seiner antiken Quellen (Aristoteles, Cicero, Augustinus, Boethius)*. Leiden: Brill.

Enders, M., Szaif, J. (2006). *Die Geschichte des philosophischen Begriffs der Wahrheit*. Berlin: de Gruyter.

Enderton, H. (1972). *Introduction to Mathematical Logic*. New York: Academic Press.

Engel, P. (1991). *The Norm of Truth: An Introduction to the Philosophy of Logic*. Toronto: University of Toronto Press.

Enayat, A., Kossak, R. (Eds.) (2003). *Nonstandard Models of Arithmetic and Set Theory*. Providence, R. I.: American Mathematical Society.

Erdmann, B. (1892). *Logik. Logische Elementarlehre*. Halle: Niemeyer; 3rd ed., Berlin: de Gruyter.

Earmann. J. (1986). *A Primer of Determinism*. Dordrecht: Kluwer.

Etchemendy, J. (1990). *The Concept of Logical Consequence*. Cambridge, Mass.: Harvard University Press.

Feferman, S. (1984) Kurt Gödel: conviction and caution. *Philosophia Naturalis*, 21, 546–562; repr. in Feferman (1998), 150–164.

Feferman, S. (1989). *The Number Systems. Foundations of Algebra and Analysis*. New York: Chelsea Publishing Company.

Feferman, S. (1998). *In the Light of Logic*. Oxford: Oxford University Press.

Feferman, S. (1999). Tarski and Gödel: between the lines. In Woleński, Köhler (1999), 53–63.

Feferman, S. (2003). Introductory note. In Gödel (2003), 43–78.

Feferman, S. (2004). Tarski's conceptual analysis of semantical notions. In Benmakhlouf (2004), 79–108.

Feferman, A, Feferman, S. (2004). *Alfred Tarski. Life and Logic*. Cambridge: Cambridge University Press.

Fenstad, J. E. (2004). Tarski, truth, and natural language. In Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E. Romanowska, A, Woleński, J. (2004), 18–26..

Field, H. (1972). Tarski's theory of truth. *Journal of Philosophy*, 69, 347–375; repr. with Postscipt in Field (2001), 3–29.

Field, H. (2001). *Truth and the Absence of Fact*. Oxford: Clarendon Press.

Flum, J. (1985). Characterizing logic. In Barwise, Feferman (1985), 77–120.

Font, J. M. (2016). *Abstract Algebraic Logic. An Introductory Textbook*. London: College Publications.

Fraenkel, A., Bar-Hillel, Y., Levy, A. (1973). *Foundations of Set Theory*. Amsterdam: North-Holland.

Frápolli, M. J. (2013). *The Nature of Truth: An Updated Approach to the Meaning of Truth Ascriptions*. Berlin: Springer.

Frege, G. (1891). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100; 25–50; Eng. tr . (by M. Black) in Frege (1984), 155–177.

Frege, G. (1892). Über Begriff und Gegenstand. *Vierteljahrsschrift für Wissenschaftliche Philosophie*, 16, 192–205; Eng. tr. in Frege (1984), 182–194.

Frege, G. (1918). Der Gedanke. *Beiträge zur Philosohie des deutschen Idealismus*, I, 58–77; Eng. tr., in Frege (1984), 351–372.

Frege, G. (1969). *Nachgelassene Schriften*. Hamburg: Felix Meiner; Eng. tr., *Posthumous Writings*. Oxford: Basil Blackwell 1979.

Frege, G. (1984). *Collected Papers on Mathematics, Logic and Philosophy*. Oxford: Blackwell.

Friedländer, P. (1954). *Platon: Seinswahrheit und Lebenswirklichkeit*. Berlin: de Gruyter; Eng. tr., *Plato*, v. 1-3, London: Routledge and Kegan Paul.

Fries, J. F. (1837). *System der Logik. Ein Handbuch zur Lehrer und Selbstgebrauch*. Heidelberg: Winter.

Frisk, H. (1936). *"Wahrheit" und "Lüge" in den indogermanischewn Sprachen. Eine morphologische Betrachtungen*. Göteborg: Elanders Boktryckeri Aktiebolag.

Gabbay, D. M., Guenther F. (Eds.) (2011), *Handbook of Philosophical Logic.* 2nd Ed., v. 11, Kluwers Academic Publishers: Dordrecht.

Gaifman, Ch. (2003). Non-standard models in a broader perspective. In Enayat, Kossak (2003), 1–22.

García-Carpintero, M., Kölbel, M. (Eds.) (2008). *Relative Truth.* Oxford: Oxford University Press.

Gardiner M. Q. (2000). *Semantic Challenges to Realism. Dummett and Putnam*. Toronto: University of Toronto Press.

Garrido, A., Wybraniec–Skardowska, U.(Eds.) (2018). *The Lvov–Warsaw School. Past and Present*. Basel: Birkhäuser.

Gaskin, R. (2013). *Language, Truth, and Literature. A Defence of Literary Humanism.* Oxford: Oxford University Press.

Gårding, L. (1977). *Encounter with Mathematics*. Springer: Berlin.

Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, 23, 121–123.

Giaro, T. 2007. *Römische Rechtswahrheiten. Ein Gedankenexperiment.* Frankfurt am Main: Vittorio Klosterman.

Gilson, E. (1955). *History of Christian Philosophy in Middle Ages*. New York: Random House.

Girard J.-Y. (1999). On the meaning of logical rules I: syntax versus semantics. In Berger, Schwichtenberg, 215–272.

Girard, J.-Y. (2011). *The Blind Spot. Lectures on Logic*. Zürich.:European Mathematical Society.

Glanzberg, M. (2003). Minimalism and paradoxes. *Synthese*, 135 (1), 12–36.

Glanzberg, M. (Ed.) (2018). *The Oxford Handbook of Truth*. Oxford University Press: Oxford.

Goldblatt, R. (1979). *Topoi. The Categorial Analysis of Logic.* Amsterdam: North-Holland.

Goble, L. (Ed.) (2001). *The Blackwell Guide to Philosophical Logic*. Oxford: Blackwell.

Goldstern, M., Judah, H. (1995). *The Incompleteness Phenomenon. A New Course in Mathematical Logic*. Wellesley: A K Peters.

Gotthardt, J. (1918). *Das Wahrheitsproblem in dem philosophischen Lebenswerk Bernard Bolzanos*. Trier: Paulinus-Drückeri.

Gómez-Torrente, M. (2004). The indefinability of truth in the "Wahrheitsbegriff". In Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E. Romanowska, A, Woleński, J. (2004), 27–38.

Gödel, K. (1930). Die Vollständigkeit der Axiome des logischen Funktionenkalkül. *Monatshefte für Mathematik und Physik*, 37, 349–360; repr. with Eng. tr. in Gödel (1986), 102–123.

Gödel, K. 1929, *Über die Vollständigkeit des Logikkalkül*, Wien Universität, repr. with Eng. tr. in Gödel 1986, 60–101.

Gödel, K. (1930). Die Vollständigkeit der Axiome des logischen Funktionenkalkül. *Monatshefte für Mathematik und Physik*, 37, 349–360; repr. with Eng. tr. in Gödel (1986), 102–123.

Gödel, K. (1930a). Einige metamathematische Resultate über Entscheidungsdefinitheit und Widerspruchsfreiheit. *Anzeiger der Akademie der Wissenschaften in Wien*, 67, 214–215; repr. with Eng. tr. in Gödel 1986, 140–143.

Gödel, K. (1931). Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38, 173–198; repr. with Eng. tr. in Gödel (1986), 145–195.

Gödel, K. (1931a). Natrag. *Erkenntnis*, 2, 149–150; repr. with Eng. tr. in Gödel 1(986), 202–204.

Gödel, K. (1932). Über Vollsändigkeit und Widerspruchsfreiheit. *Ergebnisse einer mathematischen Kolloqiums*, 3, 20–21; repr. with Eng. tr. in Gödel 1(986), 234–237.

Gödel, K. (1934). *On Undecidable Propositions of Formal Mathematical Systems* (mimeographed notes by S. C. Kleene and J. B. Rosser). Princeton: Princeton; repr. (with additions) in Gödel (1986), 346–371.

Gödel, K. (1935). Review of Carnap (1934). *Zentralblatt für Mathematik und ihre Grenzgebiete*, 11, 1; repr. with Eng. tr. in Gödel (1986), 388–389.

Gödel, K. (1944). Russell's mathematical logic. In Schillp (1944), 125–153; repr. in Gödel (1989), 119–141.

Gödel, K. (1986). *Collected Works*, v. I. Oxford: Oxford University Press.

Gödel, K. (1989). *Collected Workss*, v. II. Oxford: Oxford University Press.

Gödel, K. (1995). *Collected Works*, v. III: *Unpublished Essays and Lectures*. Oxford: Oxford University Press.

Gödel, K. (2003). *Collected Works*, v. IV: *Correspondence A–G*. Oxford: Clarendon Press. Oxford.

Gödel, K. (2003a). *Collected Works*, v. V: *Correspondence H–Z.* Clarendon Press: Oxford.

Grattan-Guiness, I. (1979). In memoriam Kurt Gödel: His 1931 correspondence with Zermelo on his incompletability theorem. *Historia Mathematica*, 6, 294–304.

Greimann, D. (Ed.) (2007). *Essays on Frege's Conception of Truth. (Grazer Philosophische Studien 75).* Amsterdam: Rodopi.

Greenough, P., Lynch, M. P. (Eds.) (2006). *Truth and Realism.* Oxford: Clarendon Press.

Gries, D., Schneider F. B. (1993). *A Logical Approach to Discrete Mathematics*. Berlin: Springer.

Grover, D. (1992). *A Prosentential Theory of Truth.* Princeton: Princeton University Press.

Gruber, M. (2016). *Alfred Tarski and the "Concept of Truth in Formalized Languages. A Running Commentary with Consideration of the Polish Original and the German Translation.* Dordrecht: Springer.

Gupta, A., Belnap, N. (1993). *The Revision Theory of Truth*. Cambridge, Mass.: The MIT Press.

Guthrie, W. K. C. (1969). *A History of Greek Philosophy*, v. 3: *The Fifth-Century Enlightenment*. Cambridge: Cambridge University Press.

Haack, S. (1978). *Philosophy of Logics.* Cambridge: Cambridge Univeristy Press.

Hàjek, P., Pudlák, P. (1998). *Metamathematics of First-Order Arithmetic.* Berlin: Springer.

Halbach, V. (2011). *Axiomatic Theories of Truth*. Cambridge: Cambridge University Press.

Halbach, V., Horsten, L. (Eds.) (2002). *Principles of Truth.* Frankfurt am Main: Dr. Hänsel-Hohenhausen.

Hallen, B. (2004). Contemporary Anglophone African philosophy: A Survey. In Wiredu (2004), 99–148.

Halldén, S. (1960). *True Love, True Humour and True Religion*. Lund: CWK Gleerup.

Hazohen, M. H. (2000). *Karl Popper. The Formative Years 1902–1045. Politics and Philosophy in Interwar Vienna*. Cambridge University Press: Cambridge.

Heidegger, M. (1940). Platons Lehre von der Wahrheit. *Geistige Überlieferung*, 2, 96–124.

Heitsch, E. (1962). Die nicht-philosophische ΑΛΗΘΕΙΑ. *Hermes. Zeitschrift für klassische Philologie*, 90, 24–33.

Hendricks, V., Neuhaus, F., Pedersen, S. A., Scheffler, U., Wansing, H. (Eds.) (2004) *First-Order Logic Revisited*. Berlin: λογος.

Henkin, L. (1949). A proof of completeness for the first-order functional calculus. *The Journal of Symbolic Logic*, 14, 159–166.

Henkin L., Addison J. Chang C. C., Craig W., Scott D., Vaught R. (Eds.) (1974). *Proceedings of the Tarski Symposium. An International Symposium Held to Honor Alfred Tarski on the Occasion of His Seventieth Birthday*. Providence: American Mathematical Society.

Herbertz, R. (1910). *Das Wahrheitsproblem in der griechischen Philosophie*. Berlin: Georg Reimer Verlag.

Hermes, H. (1952). *Mathematische Logik*. Leipzig: Teubner.

Hessen, J. (1931). *Augustinus Metaphysik der Erkenntnis*. Berlin; Ferdinand Dümmler Verlag.

Heyting, A. (Ed.) (1957). *Constructivity in Mathematics.* Amsterdam: North-Holland.

Hilbert D., Bernays P. (1939). *Grundlagen der Mathematik*, v. II. Berlin: Springer; 2nd ed., 1970.

Hill, Ch. (2002). *Thought and World. An Austere Portrayal of Truth, Reference, and Semantic Correspondence*. Cambridge: Cambridge University Press.

Hinman P. G. (2005). *Fundamentals of Mathematical Logic*. Wellesday, Mass.: A. K. Peters.

Hintikka, J. (1989). Is truth ineaffable? In Scardona (1989), 89–120; repr. in Hintikka 1997, 20–45.

Hintikka, J. (1991). *Defining Truth, The Whole Truth and Nothing but the Truth*, Reports from the Department of Philosophy, no. 2. Helsinki: University of Helsinki; repr. in Hintikka 1997, 46–103.

Hintikka, J. (1996). *The Principles of Mathematics Revisited*. Cambridge: Cambridge University Press.

Hintikka, J. (1997). *Lingua Universalis vs. Calculus Ratiocinator : An Ultimate Presupposition of Twentieth-Century Philosophy*. Dordrecht: Kluwer.

Hintikka, J. (1998). *Language, Truth and Logic in Mathematics.*. Dordrecht: Kluwer.

Hintikka, J. (Ed.) (1995). *From Dedekind to Gödel*. Dordrecht: Kluwer.

Hintikka, J., Czarnecki, T., Kijania–Placek, K., Placek, T., Rojszczak, A. (Eds.) (2003). *Philosophy and Logic. In Search of Polish Tradition. Essays in Honour of Jan Woleński on the Occasion of his 60th Birthday*. Dordrecht: Kluwer.

Hintikka, J., Puhl, K. (Eds.) (1994). *The British Tradition in 20th Century Philosophy Papers of 17th Ludwig Wittgenstein Symposium* Kirchberg am Wechsel: The Austrian Ludwig Wittgenstein Society.

Hinzen, W. (1998). *The Semantic Foundations of Anti-Realism*. Berlin: λογος.

Hipkiss, R. A. (1995). *Semantics. Defining the Discipline.* Mahwah: Erlbaum.

Hiż, H. (2004). Reexamination of Tarski's semantics. In Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E. Romanowska, A, Woleński, J. (2004), 39–48.

Hodges, W. (1985–1986). Truth in a structure. *Proceedings of Aristotelian Society*, 86, 135–151.

Hodges, W. (2004). Which languages have Tarski truth-definition. In Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E. Romanowska, A, Woleński, J. (2004), 77–92.

Hofstadter, A. (1965). *Truth and Art.* New York: Columbia University Press.

Hofstadter, A. (1965a). Truth of being. *The Journal of Philosophy*, LXII. 167–185.

Horwich, P.. (1998). *Truth.* Oxford: Clarendon Press; 2nd ed., 2000.

Horwich, P. (2005). *From a Deflationary Point of View*. Oxford: Clarendon Press.

Horwich, P. (2010). *Truth, Meaning, Reality*. Oxford: Clarendon Press.

Hoven, A. (1989). *Wege zur Wahrheit. Eine typologische Studie über Wahtheitstheorien.* Frankfurt am Main: Peter Lang.

Huby, P., Neal, G. (Eds.) (1989). *The Criterion of Truth*. Liverpool: Liverpool University Press.

Hugly, Ph., Sayward, Ch. (1996). *Intensionality and Truth. An Essay on the Philosophy of A. N. Prior*. Dordrecht: Kluwer.

Husserl, E. (1922). *Logische Untersuchungen. Band 2*, 2nd ed., Leipzig: Niemeyer; Eng. tr (by J. N. Findlay), London: Routledge 2002.

Husserl, E. (1929). *Formale und transzendentale Logik. Versuch einer Kritik der logischen Vernuft*. Halle: Niemeyer; Eng. tr., The Hague: Nijhoff 1969.

Husserl, E. (1984). *Einleitung in der Logik und Erkenntistheorie (Vortrag 1906–1907).* The Hague: Nijhoff; Eng. tr., Dordrecht: Springer 2008.

Husserl, E. (2002).*Urteilstheorie. Vorlesung 1905.* Dordrecht: Kluwer.

Ingarden, R. (1925). Essentiale Fragen. Ein Beitrag zum Problem des Wesen. *Jahrbuch für Philosophie und philosopische Forschung*, 7, 125–304.

Ingarden, R. (1949). O sądzie warunkowym (On conditional proposition). *Kwartalnik Filozoficzny*, 18, 263–308; repr. in Ingarden (1972), 271–325.

Ingarden, R. (1972). *Z teorii języka i filozoficznych podstaw logiki* (Theory of Language and Philosophical Foundations of Logic). Warszawa: Państwowe Wydawnictwo Naukowe.

Jacquette, D. Diagonalization in logic and mathematics. In Gabbay, Guenther (2011), 55–147.

Jago, M. (2018). *What is Truth*. Oxford: Oxford University Press.

James, W. (1907). *Pragmatism. A New Name for Some Old ways of Thinking. Popular Lectures on Philosophy*. London: Longmans, Green and Co.

Joachim, H. (1906).*The Nature of Truth. An Essay*. Oxford: Clarendon Press.

Joachim, H. (1907). A Reply to Mr. Moore. Mind, XVI, 410–415.

Johnson, E. L. (1992), *Focusing on Truth.* London: Routledge.

Kahn, Ch. (1973) *The Verb 'Be' in Ancient Greek*. Dordrecht: Reidel.

Kashtan, D. (2017). *An Observation about Truth (with Implications for Meaning and Language)*. Jerusalem: The Hebrew University (upublished PhD dissertation).

Kastil, A. (1947). Wahrheit und Sein. *Zeitschrift für philosophische Forschung*, I, 483–498.

Kaye R. (1991). *Models of Peano Arithmetic.* Oxford: Clarendon Press.

Ketland, J. 1999. Deflationism and Tarski's Paradise. *Mind*, 108, 69–94.

Ketland, J. 2005. Deflationism and the Gödel phenomena; Reply to Tennant. *Mind*, 114, 75–88.

Khatchadourian, H. (1961). *The Coherence Theory of Truth. A Critical Evaluation.* Beirut: Americal University.

Kirkham, R. L. (1992). *Theories of Truth. A Critical Introduction.* Cambridge, Mass: The MIT Press.

Kirkham, R. (1993). Tarski's physicalism. *Erkenntnis*, 38, 289–302.

Kittel, G. (Ed.) (1933). *Theologisches Wörterbuch zum Neuen Testament*. Stuttgart: Kohlhammer Verlag.

Knittemayer, H. (1920). *Der Terminus 'transcendental' in seiner geschichtlicher Entwicklung bis zu Kant*. Mahrburg: Phil. Doss.

Kokoszyńska, M. (1936). Syntax, Semantik und Wissenschaftlogik. In *Actes du Congès International de Philosophie Scientifique*, v. 3. Paris: Herman, 9–14; repr. in Pearce, Woleński (1988), 271–275.

Kokoszyńska, M. (1936a). .Über den Absoluten Wahrheitsbegriff und einige andere semantische Begriffe. *Erkenntnis*, 6, 143–165; repr. in Pearce, Woleński (1988), 276–292.

Kokoszyńska, M. (1936b). W sprawie względności i bezwzględności prawdy (Concerning Relativity and Absoluteness of Truth). *Przegląd Filozoficzny*, 39, 424–425.

Kokoszyńska, M. (1948). What means a „relativity" of truth". *Studia Philosophica*, III, 167–175.

Kokoszyńska, M. (1951). A refutation of the relativism of truth. *Studia Philosophica*, IV, 93–149.

Kolak, D., Symons, J. (2004). *Quantifiers, Questions and Quantum Mechanics.* Dodrecht: Springer.

Kołodziejczyk, S.T. (2006). *Granice pojęciowe metafizyki* (Conceptual Boundaries of Metaphysics), Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.

Komornicka, A. (1979). *Studia nad Pindarem and archaiczną liryką grecką. W kręgu pojęć prawdy i fałszu* (Studies on Pindar and Archaic Greek Lyric. Around the Concepts of Truth and Falsity). Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Koslow, A. (2005). Ramsey on simplicity and truth. *Metaphysica* (Special Issue), 3, 89-108.

Kossak, R. (2004). Undefinability of truth and non-standard models. In Adamowicz, Artemov, Niwiński, Orłowska, Romanowska, Woleński (2004), 115–123.

Kotarbiński, T. (1913). Zagadnienie istnienia przyszłości (The problem of the existence of future). *Przegląd Filozoficzny*, 15, 74–92.; Eng. tr, *The Polish Review*, 13(3)(1968), 7–22.

Kotarbiński, T. (1926). *Elementy logiki formalnej, teorii poznania i metodologji* (Elements of Formal Logic, Theory of Knowledge and Methodology). Warszawa: Wydawnictwo Koła Filozoficznego S. U. W. i Koła Przyrodników S. U. W.

Kotarbiński, T. (1929). *Elementy teorii poznania, logiki formalnej i metodologii nauk* (Elements of Theory of Knowledge, Formal Logic and Methodology of Science). Lwów: Ossolineum; Eng. tr. as *Gniosiology. The Scientific Approach to the Theory of Knowledge.* Oxford: Pergamon Press.

Kotarbiński, T. (1934). W sprawie pojęcia prawdy [Recenzja: A. Tarski, *Pojecie prawdy w językach nauk dedukcyjnych*] (On the concept of truth [Review of A. Tarski, *The Concept of Truth in the Languages of Deductive Sciences*]). *Przegląd Filozoficzny*, 37, 85–91; repr. in T. Kotarbiński (1993), 252–258).

Kotarbiński, T. (1993). *Ontologia*, *teoria poznania i metodologia* (Ontology, Theory of Knowledge and Methodology). Wrocław: Zakład Narodowy im. Ossolińskich.

Kotlarski, H. (2004). The incompleteness theorems after 70 years. In Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E. Romanowska, A, Woleński, J. (2004), 125–138.

Köhler, E. (2002). Gödel und Wiener Kreis. In Köhler et al. (2002), v. 1, 83–108.

Köhler, E. et al. (Eds.) (2002). *Kurt Gödel. Wahrheit & Beweisbarkeit*, Band 1.: *Dokumente und historische Analysen*; Band 2.: *Kompendium zu Werk*. Wien: öbv et htp.

Krajewski, S. (1994). Tarski's definition of truth and mathematics. In Twardowski, Woleński 1994, 16–33.

Krajewski, S. (2003). *Twierdzenie Gödla i jego interpretacje filozoficzne*. *Od mechanicyzmu do postmodernizmu* (Gödel's Theorem and Its Philosophical Interpretations. From Mechanicism to Post-Modernism). Warszawa: Wydawnictwo Instytutu Filozofii i Socjologii PAN.

Kreisel, G. (1987). Gödel's excursions into intuionistic logic. In Weingartner, Schmetterer (1987), 65–186.

Kripke, S. (1975). An outline of a theory of truth. *The Journal of Philosophy*, 72, 690–716; repr. in Kripke (2011), 75–98.

Kripke, S. (2011). *Philosophical Troubles. Collected Papers*, v. 1. New York: Oxford University Press.

Kripke, S. (2019). *Philosophical Troubles. Collected Papers*, v. 2. New York: Oxford University Press (forthcoming).

Kripke, S. (2019a). Ungroundedness in Tarskian languages. In Kripke (2019).

Krischer, T. (1965). ΕΤΥΜΟΣ und ΑΛΗΘΝΣ. *Philologus*, 109, 161–174. ***

Kumari, K. (1987). *Notion of Truth in Buddhism and Pragmatism.* New Delhi: Capital Publishing House.

Kusch, M. (1989). *Language as Calculus vs. Language as Universal Medium. A Study in Husserl, Heidegger and Gadamer*. Dordrecht: Kluwer.

Kuźniar, A., Odrowąż–Sypniewska, J. (Eds.) (2016). *Uncovering Facts and Values. Studies in Contemporary Epistemology and Political Philosophy*. Leiden: Brill/Rodoppi.

Künne, W. (1985). Wahrheit. In Martens, Schnädelbach (1985), 116–171.

Künne, W. (2002). Disquotationalist conceptions of truth. In Schantz (2002), 176–193.

Künne, W. (2003). *Conceptions of Truth.* Oxford: Clarendon Press.

Lakatos, I., Musgrave, A. (Eds.) (1968). *Proceedings of the International Colloquium in the Philosophy of Science, London, 1965*. Amsterdam: North-Holland.

Lapointe, S. (Ed.) (2015). *Themes from Ontology, Mind and Logic. Present and Past. Essays in Honour of Peter Simons*. Leiden: Brill.

Lapointe, S., Woleński, J., Marion, M., Miśkiewicz, W. (2009). *The Golden Age of Polish Philosophy. Kazimierz Twardowski's Philosophical Legacy*. Berlin: Springer.

Leblanc, H. (Ed.) (1973). *Truth, Syntax and Modality*. Amsterdam: North-Holland.

Leśniewski, S. (1913). Krytyka logicznej zasady wyłączonego środka (The critique of the logical principle of the excluded middle). *Przegląd Filozoficzny*, 16, 315–352; Eng. tr. in Leśniewski (1992), 17–85.

Leśniewski, S. (1913a). Czy prawda jest tylko wieczna czy też wieczna i odwieczna (Is all truth only true eternally or it is also true without a beginning). *Nowe Tory*, 10, 493–528; Eng. tr. in Leśniewski 1992, 86–114.

Leśniewski, S. 1929. Grundzüge eines neuen System der Grundlagen der Mathematik. *Fundamenta Mathematicae* 14, 1–81; Eng. tr. in Leśniewski 1992, 410–605.

Leśniewski, S. (1931). O podstawach matematyki VI (On the foundations of mathematics). *Przegląd Filozoficzny* 34, 142–170; Eng. tr. in Leśniewski (1992), 364–382.

Leśniewski, S. (1992). *Collected Works*. Dordrecht: Kluwer.

Levet, J. P. (1976). *Le vrai et le faux dans la pensée grecque archaïque. Étude de vocabuilare. Tome 1: Présentation générale. Le vrai et le faux dans les épopées homériques.* Paris: Les Belles Letters.

Lewis, D. (Ed.) (1956). *British Philosophy*, III. London: George Allen and Unwin.

Lindström, P. (2002). *Aspects of Incompleteness*. Natick, Mass.: A K Peters.

Linke, P. (1949). Die mehrwertige Logiken und Wahrheitsproblem. *Zeitschrift für philosophische Forschung*, III, 378–398, 530–546.

Long, Ch. R. (2011). *Aristotle on the Nature of Truth.* Cambridge: Cambridge University Press.

Lorenz, K., Mittlestrass, J. (1966). Theaitetos fliegt. Zur Theorie wahrer und falscher Sätze bei Platon (*Soph.* 251d–263d). *Archiv für Geschichte der Philosophie*, 48, 113–152.

Lowe, E. J., Rami, A. (Eds.) (2009). *Truth and Truth-Making*. Stocksfields: Acumen.

Luther, W. (1935). *"Wahrheit" und "Lüge" in ältesten Griechentum.* Borna: Robert Noske Verlag.

Luther, W. (1965). Wahrheit, Licht, Sehen und Erkennen im Sonnengleichnis von Platonis Politeia. Ein Auschnitt aus der Lichtenmetaphysik der Griechen, *Studium Generale*, 18, 479–496.

Luther, W. (1966). Wahrheit, Licht und Erkenntnis in der griechische Philosophie bis zu Demokrit. *Archiv für Begriffsgeschichte*, 10, 2–270.

Lyndon, R. C. (1966). *Notes on Logic.* New York: Van Nostrand.

Łukasiewicz, J. (1910). *O zasadzie sprzeczności u Arystotelesa. Studyum Krytyczne* (On the Principle of Contradiction in Aristotle). Kraków: Wydawnictwo Polskiej Akademii Umiejętności; Eng. tr (by H. R. Heine) *The Principle of Contradiction in Aristotle: A Critical Study.* North Charleston, SC: CreateSpace Independent Publishing 2018.

Łukasiewicz, J. (1912). O twórczości w nauce (Creative elements in science). In *Księga Pamiątkowo ku uczczeniu 250-tej rocznicy założenia Uniwersytetu Lwowskiego przez króla Jana Kazimierza* (Memorial Book on the Occasion of Establishment of Lvov University by the King Jan Kazimierz), v. II. Lwów: Nakładem Uniwersytetu Lwowskiego, 1–20; Eng. tr. in Łukasiewicz (1970), 1–15.

Łukasiewicz, J. (1913). *Die logische Grundlagen der Wahrsheilichkeitsrechnung*. Kraków: Wydawnictwo Polskiej Akademii Umiejętności; partial Eng. tr. in Łukasiewicz (1970), 16–63.

Łukasiewicz, J. (1915). O nauce (On science). In *Poradnik dla Samouków*. Warszawa: Heflich i Michalski, XV–XXXIX.

Łukasiewicz, J. (1930). Philosophische Bemerkungen zu mehrwertigen Systemem dee Aussagenkalküls . *Comptes Rendus des Sèances de la Société des Sciences et de Lettres de Varsovie, Classe III*, 51–77; Eng. tr. in Łukasiewicz (1970), 153–178.

Łukasiewicz, J (1962). O determinizmie [Mowa Rektorska z 1922 r.] (On determinism. The Rector Speech in 1922]. In Łukasiewicz (1962), 114–126; Eng. tr. (by Z. Jordan) in Łukasiewicz (1970), 110–128.

Łukasiewicz, J. (1957). *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic.* Oxford: Clarendon Press.

Łukasiewicz, J. (1962). *Z zagadnień logiki i filozofii* (Problems of Logic and Philosophy). Warszawa: Państwowe: Wydawnictwo Naukowe.

Łukasiewicz, J. (1970). *Selected Works*. Amsterdam: North-Holland.

Malinowski, G. (1994). *Many-Valued Logics.* Oxford: Oxford University Press.

Magnani, L., Bertolloti, T. (Eds.) (2017). *Springer Handbook of Model-Based Science.* Springer: Berlin.

Manzano, M. (1999). *Model Theory*.Oxford: Clarendon Press.

Mann, A. L., Sandu, G., Sevenster, M. (2011). *Independence-Friendly Logic: A Game-Theoretic Approach*. Cambridge: Cambridge University Press.

Mark, Th, (1969). *Spinoza's Theory of Truth.* New York: Columbia University Press.

Marsonet, M., Benzi, M. (Eds.) (2004). *Logic and Metaphysics. The Proceedings of Genoa Conference.* Fiesola: Name.

Martens, E., Schnädelbach, H. (Eds.) (1985). *Philosophie. Ein Grundkurs.* Reimbek bei Hamburg: Rowohlt.

Martin R. L. (Ed.) (1970). *The Paradox of Liar*. New Haven: Yale University Press.

Martin, R. L. (Ed.) (1984). *Recent Essays on Truth and the Liar Paradox.* Oxford: Clarendon Press.

Mateescu, A., Salomaa, A. (1997). Formal languages: An introduction and a synopsis. In Rosenberg, Salomaa (1997), 1–39.

Maudlin, T. (2004). *Truth and Paradox. Solving the Riddles*. Oxford: Clarendon Press.

Mauthner, F. (1902). *Zur Grammatik und Logik.* Stuttgart: Cotta.

McCloskey, H. J. (1969). *Meta-Ethics and Normative Ethics*. Berlin: Springer.
McGee, V. (1991). *Truth, Vagueness and Paradox.* Indianapolis: Hackett Publishing Company.
McGinn C. (1980). Truth and use. In Platts (1980), 19–40.
McGrath, M. (2000). *Between Deflationism and Correpondence Theory*. New York: Garland.
Meier, H. (1926). *Wahrheit und Wirklichkeit*. Tübingen: Mohr.
Mendelson, E. (1997). *Introduction to Mathematical Logic*. London: Champan & Hall. London.
Menger, K. (1994). *Reminiscences of the Vienna Circle and the Mathematical Colloqium*. Dordrecht: Kluwer.
Merricks, T. (2007). *Truth and Ontology*. Oxford: Clarendon Press.
Michel, D. (1968). Ämät. Untersuchungen über "Wahrheit" in Hebräischen". *Archiv für Begriffsgeschichte*, 12, 30–57.
Moore, G. E. (1899). The nature of judgement. *Mind*, 5, 176–193; repr. in Moore 1993, 1–19.
Moore, G. E. (1907). Mr. Joachim's *Nature of Truth*. Mind, XV, 229–235.
Moore, G. E. (1953). *Some Main Problems of Philosophy*.London: George Allen & Unwin.
Moore, G. (1993). *Selected Writings.* London: Routledge.
Moreno, L. F. (1992). *Wahrheit und Korrespondenz bei Tarski. Eine Untersuchung der Wahrheitstheorie Tarskis als Korrepondenztheorie der Wahrheit*. Würzburg: Köningshausen & Neumann.
Morgan, M., Morrison, M. (1999). *Models as Mediators. Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.
Morris, Ch. (1938). *Foundations of the Theory of Signs.* Chicago: University of Chicago Press.
Morton, A., Stich, S. P. (Eds.) (1996). *Benacerraf and His Critics*. Oxford: Blackwell.
Mostowski, A. (1948). *Logika matematyczna* (Mathematical Logic. Warszawa: Monografie Matematyczne.
Mostowski, A. (1952). *Sentences Undecidable in Formalized Arithmetic*. Amsterdam: North-Holland.
Mostowski, A. (1957). On various degrees of constructivism. In Heyting (1957), 178–194, repr. in Mostowski (1979), v. II, 359–375.
Mostowski, A. (1965). *Thirty Years of Foundational Studies. Lectures on the Development of Mathematical Logic and the Study of the Foundations of Mathematics in the Years 1930–1963.* Helsinki: Societas Philosophica Fennica; repr. in Mostowski (1979), v. I, 1–176.
Mostowski, A. (1967). Alfred Tarski. In Edwards (1967), v. 8, 77–81.
Mostowski, A. (1979). *Foundational Studies. Selected Works*, v. I–II. Amsterdam: North-Holland.
Mulligan, K., Simons, P. M., Smith, B. (1984). Truth-makers. *Philosophy and Phenomenological Research* 44, 287–321; repr. in Lowe, Rami (2009), 59–86.
Murawski, R. (1998). Undefinability of truth. The problem of priority: Tarski vs. Gödel. *History and Philosophy of Logic*, 19, 153–160.
Murawski, R. (1999). *Recursive Functions and Metamathematics. Problems of Completeness and Decidability*. Dordrecht: Kluwer.
Murawski, R. (1999a). Undefinability vs. definability of satisfaction and truth. In Woleński, Köhler (1999), 203–215.
Murawski, R., Woleński, J. (2008). Tarski and his precedessors on truth. In Patterson (2008), 21–43.
Mycielski, J. (2004). On the tension between Tarski's nominalism and model theory (Definitions for a mathemathical model of knowledge). In Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E. Romanowska, A, Woleński, J. (2004), 215–224.
Naess A. (1938). *„Truth" as Conceived by Those Who Are Not Professional Philosophers"*. Oslo: I Kommisjon Hos Jacob Dybward.
Naess, A. (1953). *An Empirical Study of the Expression „True", „Perfectly Certain" and „Extremely Probable"*. Oslo: I Kommisjon Hos Jacob Dybward.
Neale, S. (2001). *Facing Facts*. Oxford: Clarendon Press.
Nesher, D. (2002). *On Truth and the Representation of Reality. A Collection of Inquiries from a Pragmatist Point of View*. Landham: University Press of America.

Newman, A. (2002). *The Correspondence Theory of Truth. An Essay on the Metaphysics of Predication.* Cambridge: Cambridge: University Press.

Niiniluoto, I. (1987). *Truthlikeness.* Dordrecht: Kluwer.

Niiniluoto, I. (1994). Defending Tarski against his critics. In Twardowski, Woleński (1994), 48–68.

Niiniluoto, I. (1999). Theories of truth: Vienna, Berlin and Warsaw. In Woleński, Köhler (1999), 17–26.

Niiniluoto, I. (2004). Tarski's definitions and truth-makers. In Adamowicz, Z., Artemov, S., Niwiński, D., Orłowska, E. Romanowska, A, Woleński, J. (2004), 57–76.

Niiniluoto, I., Sintonen, M., Woleński, J. (Eds.) (2004). *Handbook of Epistemology*. Dodrecht: Kluwer.

Novak, Z., Simonyi, A. (Eds.) (2011). *Truth, Reference and Realism.* Budapest: CEU Press.

O,'Connor, D. J. O. (1975). *The Correspondence Theory of Truth*. London: Hutchinson.

Ogden, C. K., Richards, I. A. (1923). *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism*. London: Routledge and Kegan Paul.

Olsson, E. J. (2005). *Against Coherence. Truth, Probability and Justitication*. Oxford: Clarendon Press.

Pantsar, M. (2009). *Truth, Proof and Gödelian Arguments. A Defence of Tarskian Truth in Mathematics*. Helsinki: University of Helsinki.

Pap, A. (1954). Propositions, sentences and the semantic concept of truth, *Theoria*, XX, 23–35.

Pardey, U. (2012). *Freege on Absolute and Relative Truth*. New York: Palgrave Macmillan.

Parsons, Ch., Schwartz, R. (Eds.). *How Many Questions? Essays in Honor of Sidney Morgenbesser*. Indianapolis: Hackett Publishing Company.

Parsons, D. (2016). *Theories of Intensionality. A Critical Survey*. Dordrecht: Springer.

Pataut, F. (Ed.) (2016). *Truth, Objects, Infinity. New Perspectives on the Philosophy of Paul Benacerraf.* Dordrecht: Springer.

Patterson, D. E. (2012). *Alfred Tarski, Philosophy of Language and Logic.* Basingstoke: Palgrave Macmillan.

Patterson, D. E. (2006). Tarski and the necessity reading of the Convention T. *Synthese*, 151(1), 1–32.

Patterson, D. E. (Ed.) (2008). *New Essays on Tarski and Philosophy*. Oxford: Oxford University Press.

Patterson, D. M. (1999). *Law and Truth.* Oxford: Oxford University Press.

Pearce, D., Woleński, J. (Eds.) (1988). *Logische Rationalismus. Philosophische Schriften der Lemberg–Warschauer Schule*. Frankfurt a. M.: Athenäum.

Pedersen, N. J., Wright, C. D. (2013). *Truth and Pluralism.* Oxford: Oxford University Press.

Pelettier, F. J. (1990). *Parmenides*, *Plato and the Semantics of Not-Being*. Chicago: The University of Chicago Press.

Peregrin, J. (Ed.) (1999). *The Nature of Truth (If Any).* Dordrecht: Kluwer.

Perler, D. (1992). *Der Propositionale Wahrheitsbegriff im 14. Jahrhundert*. Berlin: de Gruyter.

Perzanowski, J. (1996). The way of truth. In Poli and Simons (1996), 61–130.

Peters, F. E. (1967). *Greek Philosophical Terms. A Historical Lexicon*, New York: New York University Press.

Platts, M. (Ed.) (1980). *Reference, Truth and Reality*. London; Routledge and Kegan Paul.

Pogorzelski, W. A. (1994). *Notions and Theorems of Elementary Formal Logic.* Białystok: Białystok University Press.

Poizat, B. (2000). *A Course in Model Theory. An Introduction to Contemporary Mathematical Logic*. Berlin: Springer.

Poli, R., Simons, P. (Eds.)(1996). *Formal Ontology*. Dordrecht: Kluwer.

Popper, K. (1955). A note on Tarski's definition of truth. *Mind*, 64, 388–391; repr. in Popper (1972), 335–340.

Popper, K. (1972). *Objective Knowledge. An Evolutionary Approach.* Oxford: Clarendon Press; repr. with additions, 1974.

Prawitz, D., Skyrms, B., Westerståhl, D. (Eds.) (1994). *Logic, Methodology and Philosophy of Science IX*. Amsterdam: Elsevier.

Priest, G. (2006). *Doubt Truth to Be a Liar*. Oxford: Clarendon Press.

Priest, G. (2008). *An Introduction to Non-Classical Logic.* Cambridge: Cambridge: University Press.

Prior, A. N. (1971). *Objects of Thought*. Oxford: Clarendon Press.

Pritzl, K. (Ed.) (2010). *Truth. Studies in a Robust Presence.* Washington, D. C.: The Catholic University of America Press.

Przełęcki, M. (1969). *The Logic of Empirical Theories.* London: Routledge and Kegan Paul.

Przełęcki, M. (1974). Some philosophical consequences of the semantic definition of truth. *Dialectic and Humanism*, I, 117–128.

Przełęcki, M. (1974a). On model-theoretic approach to interpretation of empirical theories. *Synthese*, 26, 401–406.

Puntel, L. B. (1990). *Grundlagen einer Thorie der Wahrheit.* Berlin: de Gruyter.

Putnam, H. (1975). *Mind, Language and Reality. Philosophical Papers*, v. 2. Cambridge: Camrbidge University Press.

Putnam, H. (1975a). Do true assertions correspond to reality? In Putnam 1975, 70–84.

Putnam H. (1980). Models and reality. *The Journal of Symbolic Logic*, 464–482.

Putnam, H. (1983). On truth. In Cauman, Levi, Parsons, Schwartz 1983, 35–56; repr. in Putnam (1994), 315–329.

Putnam, H. (1985–1986). On comparison of something with something else. *New Literary History*, 17, 61–79; repr. in Putnam 1994, 330–350.

Putnam, H. (1994). *World and Life*. Cambridge, Mass.: Harvard University Press.

Quell, G. (1933). Der at.liche Begriff "ämät". In Kittel 1933, 233–237.

Quine, V. v. O. (1940). *Mathematical Logic.* Cambridge, Mass.: Harvard University Press.

Quine, V. v. O. (1953). *From a Logical Point of View.* Cambridge, Mass.: Harvard University Press.

Quine, V. V. O. (1953a). Notes on the theory of reference. In Quine (1953), 130–138.

Quine, W. v. O. (1970). *Philosophy of Logic*. Englewood Clifts: Prentice Hall.

Quine, W. v. O. (1987). *Quiddities. An Intermittently Philosophical Dictionary*. Cambridge, Mass.: Harvard University Press.

Quine, W. v. O. (1990). Lectures on Carnap. In Quine, Carnap (1990), 407–103.

Quine, V. v. O., Carnap, R. (1990). *Dear Carnap, Dear Van: The Quine–Carnap Correspondence and Related Work*. Los Angeles: University of California Press.

Rami, A. (2009). *Wahrheit und Deflation. Eine kritische Untersuchung deflationärer Wahrheitskonzeptionen.* Paderborn: Mentis.

Rami, A. (2009a). Introduction: Truth and truth-making. In Loewe, Rami (2009), 1–36.

Ramsey, F. P. (1925). The foundations of mathematics. *Proceedings of the London Mathematical Society* (Series 2), 25, 338–384; repr. in Ramsey 1931, 1–61 and Ramsey (1978), 152–212.

Ramsey, F. P. (1927). Facts and propositions. *Proceeding of Aristotelian Society. Supplementary Volume*, VII, 153–170; repr. in Ramsey (1931), 138–155 and Ramsey (1978), 40–57.

Ramsey, F. P. (1931). *The Foundations of Mathematics*. London: Routledge and Kegan Paul; 2nd ed., *Foundations*. *Essays in Philosophy*, *Logic*, *Mathematics and Economics*. London: Routledge and Kegan Paul 1978.

Ramsey, F. P. (1991). *On Truth. Original Manuscript Materials (1927–1929) from the Ramsey Collection at the University of Pittsburgh.* Dordrecht: Kluwer.

Ramsey, F. P. (2007). Truth and simplicity. *British Journal of Philosophy of Science*, 58, 1–7.

Rasmussen, Y. (2004). *Defending the Correspondence Theory of Truth.* Cambridge: Cambridge University Press.

Rauzy, J.-B. (2001). *La doctrine Leibnizienne de la vérité. Aspectes logiques et ontologiques*. Paris: Vrin.

Ray, G. (2018). Tarski on the concept of truth. In Glanzberg (2018), 695–719.

Reboul, A. (Ed.) (2014). *Mind, Values, and Metaphysics. Philosophical Essays in Honor of Kevin Mulligan*, v. 1. Dordrecht: Springer.

Rescher, N. (1969). *Many-Valued Logic.* New York: McGraw-Hill.

Rescher, N. (1973). *The Coherence Theory of Truth*. Oxford: Clarendon Press.

Rescher, N. (1985). Truth as ideal coherence. *The Review of Metaphysics*, 38, 795–806.

Rittter, J., Gründer, K.-H., Gabriel, G. (Eds.) (1971–2007). *Historischer Wörterbuch der Philosophie*,. I–XIII. Basel: Benno Schwabe.

Rojszczak, A. (2005). *From the Act of Judging to the Sentence. The Problem of Truth Bearers from Bolzano to Tarski*. Dordrecht: Springer.

Rojszczak, A., Cachro, J., Kurczewski, G. (Eds.) (2003). *Dimensions of Logic and Science. Selected Contributed Papers from the 11th International Congress of Logic, Methodology, and Philosophy of Science*. Kraków, 1999. Dordrecht: Kluwer.

Rosenberg, G, , Salomaa, A. (Eds.) (1997). *Handbook of Formal Languages*, v. 1: *Word, Language, Grammar*. Berlin: Springer.

Rosser, J. B. (1936). Extensions of some theorems of Gödel and Church. *The Journal of Symbolic Logic*, 1, 87–91; repr. in Davis (1965), 231–235).

Russell, B. (1903). *The Principles of Mathematics*. Cambridge: Cambridge University Press. Cambridge.

Russell, B. (1904). Meinong's theory of complexes and assumptions" (III). *Mind*, XIII, 509–524.; repr. in Russell 1973, 54-69.

Russell, B. (1906). The nature of truth. *Mind*, XV, 528–533.

Russell, B. (1906a). What is truth? *Independent Review*, 9, 349–353.

Russell, B. (1907). On the nature of truth. *Proceedings of Aristotelian Society* (New Series), VII, 28–49.

Russell, B. (1910). *Philosophical Essays*. London: Longmans.

Russell, B. (1912). *The Problems of Philosophy*. London: Wiliam and Norgate.

Russell, B. (1940). *An Inquiry into Meaning and Truth.* London: George Unwin and Allen.

Russell, B. (1973). *Essays in Analysis*. New York: George Braziller.

Russell, B. (1984). *The Theory of Knowledge. The 1913 Manuscript*. London: George Allen and Unwin.

Rüstow, A (1910). *Der Lügner. Theorie,Geschichte und Auflösumg*. Leipzig: Teubner Verlag.

Ryle, G. (1953). Ordinary language. *Philosophical Review* XLII, 167–186; repr. in Chappell (1964), 24–40.

Scardona N. (Ed.) (1989). *Les Formes Actuelles du Vrai: Entretiens de Palermo*. Palermo: Enhiridion.

Schantz, R. (Ed.) (2002). *What is Truth?* Berlin: de Gruyter.

Scaglia, B. (2011). *The Epistemology of Truth.: Veritas, Logic and Interpretation.* Memphis: Books LLC.

Scheffer, Th. (1993). *Kants Kriterium der Wahrheit. Anschauungsformen und Kategorien a priori in der "Kritik der reinen Vernuft"*. Berlin: de Gruyter.

Schillp, P. A. (Ed.) (1944). *The Philosophy of Bertrand Russell*. La Salle: Open Court.

Schilpp, P. A. (Ed.) (1963). *The Philosophy of Rudolf Carnap*. La Salle: Open Court.

Schlick, M. (1918). *Allgemeine Erkenntnislehre*. Springer, Berlin; Eng. tr., *General Theory of Knowledge*. Wien: Springer.

Schlick, M. (1931). Das Kausalität in den gegenwärtigen Physik. *Naturwisseschaften*, 19, 145–162; Eng. tr. in Schlick (1979), 176–209.

Schlick, M. (1979). *Collected Papers*, v. II. Dordrecht: Reidel.

Scholz, H. (1950). *Vorlesungen über Grundzüge der mathematischen Logik.* Münster: Aschendoffsche Verlagsbuchhandlung.

Schröter, J. (1996). *Zur Meta-theorie der Physik.* Berlin: de Gruyter.

Schulemann, G. (1929). *Die Lehre von der Transzendentalien in der scholastischen Philosophie*. Leipzig: Teubner.

Schulz, G. (1993). *Veritas est adequatio intellectus et rei. Untersuchungen zur Wahrheitslehre des Thomas von Aquin & zur Kritik Kants an einen überrlieferten Wahrheitsbegriff*. Leiden: Brill

Senner, W. (1995). Zur Definition der Wahrheit bei Albertus Magnus. In Eggensperger, Engel (1995), 11–48.

Shapiro, S. (1991). *Foundations without Foundationalism*. Oxford: Oxford University Press.

Shapiro, S. (Ed.) (1996). *The Limits of Logic. Higher-order Logic and the Löwenheim–Skolem Theorem*. Aldershot: Dartmouth.

Shapiro, S. (1996a). Introduction. In Shapiro (1996), XI–XXII.

Shapiro, S. (2014). *Varieties of Logic.* Oxford: Oxford University Press.

Sher, G. (1999). On the possibility of of a substantive theory of truth. *Synthese*, 117, 133–172.

Shope, R. K. (1983). *The Analysis of Knowledge. A Decade of Research.* Princeton: Princeton University Press.

Simons, P. (1992). *Philosophy and Logic in Central Europe from Bolzano to Tarski. Selected Essays*. The Hague: Nijhoff.

Simons P. (2003). Absolute truth in a changing world. In Hintikka, Czarnecki, Kijania–Placek, Placek, Rojszczak (2003), 37–54.

Shramko, Y., Wansing, H. (2011). *Truth and Falsehood. An Inquiry into Generalized Logical Values.* Berlin: Springer.

Sierpiński, W. (1965). *Cardinal and Ordinal Numbers.* Warszawa: Państwowe Wydawnictwo Naukowe.

Simmons, K. (1993). *Universality and the Liar. An Essay on Truth and the Diagonal Argument*. Cambridge: Cambridge University Press.

Simpson, S. G. (1999). *Subsystems of Second-Order Arithmetic*. Berlin: Springer.

Sinnott-Armstrong, W., Raffman, D, Asher, N. (Eds.) (1995). *Modality, Morality and Belief. Essays in Honor of Ruth Barcan Marcus*. Cambridge: Cambridge University Press.

Sirakov, B, Ney de Souza, P., Viana, M. (Eds.) (2019). *Proceedings of International Congress of Mathematicians. Rio de Janeiro 2018,* v. 3. Singapore: World Scientific.

Smorynski, C. (1985). *Self-Reference and Modal Logic.* Berlin: Springer.

Smullyan, R. M. (1992). *Gödel's Incompleteness Theorem.* Oxford: Oxford University Press.

Smullyan, R. M. (1994). *Diagonalization and Self-Reference. Oxford:* Oxford University Press.

Smullyan, R. M. (2001). Gödel's incompleteness theorems. In Goble (2001), 72–89.

Snell. B. (1975). ΑΛΗΘΕΙΑ. *Würzburger Zeitschsrift für die Altertumwissenschaft. Neue Folge*, 1, 9–17.

Snell, B. (1978). *Der Weg zum Denken und zur Wahrheit. Studien zur frügrechischen Sprache.* Göttingen: Vandenhoeck & Ruprecht.

Soames, S. (1995). T-sentences. In Sinnott-Armstrong, Raffman, Asher (1995), 250–270.

Soames, S. (1999). *Understanding Truth.* Oxford: Oxford University Press.

Sprigge, T. L. S. (1993). *James & Bradley. American Truth & British Reality*. La Salle: Open Court.

Srzednicki, J. (1965). *Franz Brentano's Analysis of Truth.* The Hague: Nijhoff.

Stegmüller, W. (1957). *Das Wahrheitsbegriff und die Idee der Semantik*. Wien: Springer.

Stelmach, J., Brożek, B, Kwiatek, Ł. (Eds.) (2016). *The Normative Mind*. Kraków: Copernicus Center.

Suppes, P. (1988). Philosophical implications of Tarski's work. *The Journal of Symbolic Logic*, 53, 80–91.

Suppes P. (2002). *Representation and Invariance of Scientific Structures*. Stanford: CSLI Publications.

Surma, S. J. (1981). The growth of logic out of the foundational research in the foundations of mathematics. In Agazzi (1981), 15–33.

Suszko, R. (1968). Formal logic and the development of knowledge. In Lakatos (1968), 210–222.

Szaif, J. (1996). *Platos Begriff der Wahrheit*. Freiburg: Karl Alber Verlag.

Szaniawski, K. (Ed.) (1989). *The Vienna Circle and the Lvov–Warsaw School*. The Hague: Nijhoff.

Tarski, A. (1930). Über einige fundamentale Begriffe der Metamathematik. *Comptes Rendus des séances de la Société des Sciences et de Lettres de Varsovie*, 23, 22–29; Eng. tr. in Tarski (1956), 30–38.

Tarski, A. (1930a). Über definierbare Zahlenmengen. *Annales de Société Polonais Mathématique*, 9, 206–207; repr. in Tarski 1986, v. 4, 560–561.

Tarski, A. (1930–1931). O pojęciu prawdy w odniesieniu do sformalizowanych nauk dedukcyjnych (On the Concept of Truth in Reference to Formalized Deductive Sciences). *Ruch Filozoficzny*, 12, 210–211; repr. in Tarski 1986, v. 4, 559.

Tarski, A. (1930–1931a). O pewnym systemie logiki matematycznej i wynikających z niego zagadnieniach metodologicznych (On Certain System of Mathematical Logic and the Resulting Methological Issues). *Ruch Filozoficzny*, 12, 222; Eng. tr . (by J. Tarski) in Woleński (2004c), 123.

Tarski A. (1931). Sur les ensembles définissables de nombres réels. I. *Fundamenta Mathematicae*, XVII, ss. 210–239; Eng. tr. in Tarski 1956, 110–142.

Tarski, A. (1932). Der Wahrheitsbegriff in der Sprachen der deduktiven Wissenschaften. *Akademie der Wissenschaften in Wien. Matematisch-Wissenschaftliche Anzeiger*, 69, 23–25; repr. In Tarski 1986, v. 1, 613–617.

Tarski, A. (1933). *Pojęcie prawdy w językach nauk dedukcyjnych* (The Concept of Truth in Languages of Deductive Sciences). Warszawa: Towarzystwo Naukowe Warszawskie; Germ. tr. (with additions) as Tarski (1935).

Tarski, A. (1933a). Einige Betrachtungen über die Begriffe ω-Widerspruchsfreiheit und der ω-Vollständigkeit. *Monatshefte für Mathematik un Physik*, 40, 97–112; repr. in Tarski 1956, 279–295.

Tarski, A. (1935). Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica*, 1, 261–405; repr. in Tarski (1986), v. II, 51–198. Eng. tr. (The Concept of Truth in Formalized Languages), in Tarski (1956), 152–278.

Tarski A. (1936). Grundlegung der wissenschaftliche Semantik. In *Actes du Congès International de Philosophie Scientifique*, v. 3. Paris: Herman, 1–8; Eng. tr. in Tarski (1956), 401–408.

Tarski A. (1936a) Über den Begriff der logischen Folgerung. In *Actes du Congrès International de Philosophie Scientifique*, v. 7. Paris: Herman, 1–11; Eng. tr. in Tarski (1956), 409–420.

Tarski (1936b). [Głos w dyskusji nad referatem M. Kokoszyńskiej] (Contribution to the dicscussion on M. Kokoszyńska's talk). *Przegląd Filozoficzny*, 39, 425; repr. in Tarski 1986, v. 4, 701.

Tarski, A. (1939). On undecidable sentences in enlarged systems of logic and the concept of ruth. *The Journal of Symbolic Logic*, 4, no 3, 105–112; repr. in Tarski (1986), v. 2, 559–570.

Tarski, A. (1944). Semantic theory of truth and the foundations of semantics. *Philosophy and Phenomenological Research* 4; repr. in Tarski (1986), v. 2, 661–697.

Tarski, A. (1954). [Contribution to the discussion of P. Bernays, *Zur Beurteilung der Situation in der beweistheoretischen Forschung*]. *Revue Internationale de Philosophie*, 8, 16–20; repr. in Tarski (1986), v. 4, 713–714.

Tarski, A. (1954a). [Remarks of Alfred Tarski]. *Revue Internationale de Philosophie*, 27–28, 18–19; repr. in Tarski (1986), v. 4, 713.

Tarski, A. (1956). *Logic*, *Semantics*, *Metamathematics*. *Papers of 1923 to 1938* (tr. by J. H. Woodger). Oxford: Clarendon Press; 2nd ed., Indianapolis: Hackett Publishing Company 1983.

Tarski, A. (1969). Truth and proof. *Scientific American*, 220(6), 63–77; repr. in Tarski (1986), v. 4, 399–423.

Tarski, A. (1986). *Collected Papers*, v. 1–4. Basel: Birkhäuser; 2nd ed. 2018,

Tarski, A. (1986a). What are logical notions? *History and Philosophy of Logic*, 7, 143–154.

Tarski, A. (1992). Drei Briefe an Otto Neurath [25. IV. 1930, 10. VI. 1936, 7. IX. 1936], together with Eng. tr. by J. Tarski, 1–29. *Grazer Philosophische Studien*, 42, 1–29.

Tarski, A. (2001). *Pisma logiczno-semantyczne* (Logico-Semantic Writings). v. 2: *Metalogika* (Metalogic). Warszawa: Wydawnictwo Naukowe PWN.

Tarski, A. (2001a). Odczyt Alfreda Tarskiego na *Konferencji o Problemach Matematyki w Princeton, 17 grudnia 1946* (Alfred Tarski's Lecture at the Conference *On Problems of Mathematics,* Princeton, *November 17, 1946*). In Tarski 2001, 396–413.

Tarski, A., Givant, S. (1987). *A Formalization of Set Theory without Variables.* Providence, L. I.: American Mathematical Society.

Tarski, A., Mostowski, A., Robinson, R. M. (1953). *Undecidable Theories.* Amsterdam: Norh-Holland.

Tarski, A., Vaught, R. (1957). Arithmetical extensions of relational systems. *Compositio Mathematica*, 13, 81–102; repr. in Tarski 1986, v. 4, 651–682.

Taylor, B. (2006). *Truth and Realism.* Oxford : Oxford University Press.

Tennant, N. (1987). *Anti-Realism and Logic. Truth as Eternal.* Oxford : Clarandon Press.

Tennant, N. (1997). *The Taming of the True.* Oxford : Clarendon Press.

Tennant, N. (2002). Deflationism and the Gödel phenomena. *Mind*, 111, 551–582.

Tennant, N. (2005). Deflationism and the Gödel phenomena: reply to Ketland. *Mind*, 114, 89–96.

Tuboly, A. T. (2017). From 'Syntax to 'Sematnik' – Carnap's inferentialism and its prospects (unupblished manuscript).

Turner, R. (1991). *Truth and Modality for Knowledge Representation*. London: Pitman.

Turner, R. (1999). Logics of truth. *Notre Dame Journal of Formal Logic*, 31, 308–329.

Twardowski, B., Woleński, J. (Eds.) (1994). *60 Years of Tarski's Definition of Truth*. Kraków: Phileb.

Twardowski, K. (1900). O tzw. prawdach względnych (On so-called relative truths). In *Księga Pamiątkowa Uniweysytetu Lwowskiego ku uczczeniu pięćsetnej rocznicy Fundacji Jagiellońskiej Uniwersytetu Krakowskiego* (Memorial Book of Lvov University on the Occasion of 500th Anniversary of Jagiellonian Fund – the University of Kraków). Lwów: Nakładem Uniwersytetu Lwowskiego, 64–93; Eng. tr. in Twardowski 1999, 147–169.

Twardowski, K. (1912). O czynnościach i wytworach. Kilka uwag z pogranicza psychologii, gramatyki i logiki (Actions and products. Remarks from the borderline of psychology, grammar and logic). In *Księga Pamiątkowo ku uczczeniu 250-tej rocznicy założenia Uniwersytetu Lwowskiego przez króla Jana Kazimierza* (Memorial Book on the Occasion of Establishment of Lvov University by the King Jan Kazimierz), v. II. Lwów: Nakładem Uniwersytetu Lwowskiego, 1–33; Eng. tr. in Twardowski 1999, 103–132.

Twardowski, K. (1975). Wykłady z teorii poznania w r. A. 1924–25 (Theory of knowledge. Lecture course 1924–25). *Archiwum Historii Filozofii i Myśli Społecznej*, 22, 241–299; Eng. tr. in Twardowski (1999), 181–239.

Twardowski, K. (1997). Dzienniki (Daily Diaries), v. 1–2. Toruń: Adam Marszałek.

Twardowski, K. (1999). *On Actions, Products and Other Topics in Philosophy* (tr. by A. Szylewicz). Amsterdam: Rodopi.

Udemadu, K. (1995). *Einwände gegen Alfrer tarski semantische Theorie der Wahrheit*. Norderstedt: Grin.

Ulatowski, J. (2016). Ordinary truth in Tarski and Næss. In Kuźniar, Odrowąż–Sypniewska (2016), 67–89.

Ulatowski, J. (2017). *Commonsense Pluralism about Truth An Empirical Defence*. Basinstoke: Palgrave Macmillan.

Van Atten, M. (2015). *Essays on Gödel's Reception of Leibniz, Husserl, and Brouwer.* Dordrecht: Springer.

Van De Mieroop, M. (2015). *Philosophy before the Greeks: The Pursuit of Truth in Ancient Babylonia*. Hoboken, NJ.: Wiley-Blackwell.

Van Fraassen, B. (1971). *Formal Semantics and Logic.* New York: Macmillan.

Van Heijenoort, J. (1967). Logic as calculus and logic as language. *Synthese*, 17, 324–330; repr. in Van Heijenoort (1985), 11–16.

Van Heijenoort, J. (1985). *Selected Essays*. Napoli: Bibliopolis.

Van Zantwijk, T., Gabriel G. (2001). Urteil. In Ritter, Gründer, Gabriel (1971–2007), v. XI, 430–458.

Vaught, R. (1974). Model theory before 1945. In Henkin, Addison, Chang, Craig, Scott, Vaught (Eds.) (1974), 153–172.

Vaught, R. (1986). Alfred Tarski's works in model theory. *Journal of Symbolic Logic*, 51, 861–882.

Vänäänen, J. (1985). Set-theoretical definability of logics. In Barwise, Feferman (1985), 599–643.

Vänäänen, J. (2001). Second-order Logic and the Foundations of Mathematics. *Bulletin of Symbolic Logic*, 7, 504–520.

Vernant, J.–P. (1983). *Myth and Thought among Greeks*. London: Methuen.

Vinci, Th. C. (1998). *Cartesian Truth*. New York: Oxford University Press.

Vision, G. (1988). *Modern Anti-Realism and Manufactured Truth*. London: Routledge.

Vision, G. (2004). *Veritas: The Correspondence Theory and Its Critics*. Cambridge, Mass.: The MIT Press.

Visser, A. (2011). Semantics and the Liar Paradox. In Gabbay, Guenther (2011), 149–240.

Von Neumann, J. (1966). *Theory of Self-Reproducing Automata*. Urbana: University of Illinois Press.

Von Plato, J. (2019). In search of the source of incompleteness. In Sirakov, Ney de Souza, Viana (2019), 4043–4060.

Von Wright, G. H. (1996). *Six Essays in Philosophical Logic*. Helsinki: Societas Philosophica Fennica.

Von Wright, G. H. (1996a). Truth logics. In Von Wright (1996), 71–91.

Vroom, H. M. (1989). *Religions and Truth: Philosophical Reflections*. Amsterdam: Rodopi

Vuillemin, J. (1996). *Necessity or Contingency. The Master Argument.* Stanford: CSLI Publications.

Waismann, F. (1956). How I see philosophy? In Lewis (1956), 447–490; repr., in Waismann 1968, 1–38.

Waismann, F. (1965). *The Principles of Linguistic Philosophy*. London: Macmillan,

Waismann, F. (1968). *How I See Philosophy*.London: Macmillan.

Waismann, F. (1976). *Logik, Sprache, Philosophie*. Stuttgart: Reclam.

Walker, R. C. (1989). *The Coherence Theory of Truth. Realism, Anti-Realism, Idealism*. London: Routledge.

Wang, H. (1986). *Beyond Analytic Philosophy. Doing Justice to What we Know*. Cambridge, Mass.: The MIT Press.

Wang, H. (1987). *Reflections on Kurt Gödel*. Cambridge, Mass.: The MIT Press.

Wang, H. (1996). *A Logical Journey. From Gödel to Philosophy*. Cambridge: Mass: The MIT Press.

Weaver, N. (2015). *Truth & Assertibility.* Singapore: World Scientific.

Weingartner, P. (2000). *Basic Questions on Truth.* Dordrecht: Kluwer.

Weingartner, P., Schmetterer, L. (Eds.) (1987). *Gödel Remembered.* Napoli: Bibliopolis.

Wentscher, E. (1941). *Relative oder absolute Wahrheit? Eine Studie*. München: Reinhardt.

Wessel, H. (1999). *Logik und Philosophie*. Berlin: λογος.

Westerståhl, D. (1976). *Some Philosophical Aspects of Abstract Model Theory*. Göteborg: Institutionen för Filosofi, Göteborgs Universitet.

Wiesner, J. (1996). *Parmenides. Der Beginn der Aletheia. Untersuchungen zu B2–B3–B6.* Berlin: de Gryuter.

Williams, B. (2004). *Truth and Truthfulness. An Essays in Genealogy.* Princeton: Princeton University Press.

Williams, C. J. F. (1976). *What is Truth?* Cambridge: Cambridge University Press.

Williamson, T. (2013). *Modal Logic as Metaphysics.* Oxford: Oxford University Press.

Wiredu, K. (Ed.) (2004). *A Companion to African Philosophy.* Oxford: Blackwell.

Whitehead, A. N., Russell, B. (1910–1913). *Principia Mathematica.* Cambridge: Cambridge University Press: Cambridge; 2nd ed., 1925.

Wittgenstein, L. (1922) *Tractatus Logico-Philosophicus*, tr. by Ch. Ogden. London: Routledge and Kegan Paul.

Woleński, J. (1989). *Logic and Philosophy in the Lvov–Warsaw School*. Dordrecht: Kluwer.
Woleński, J. (1989a). Brentano's criticism of the correspondence theory of truth and Tarski's semantic theory. *Topoi*, 6(1989), 105–110; repr. in Woleński (1999), 103–109.
Woleński, J. (1990a). Kotarbiński, many-valued logic and truth. In Woleński (1990), 190–197; repr. in Woleński (1999), 115–120.
Woleński, J. (1993). Two concepts of correspondence. *From the Logical Point of View*, II(3), 42–55.
Woleński, J. (1994). Controversy over the concept of correspondence (Bradley, Joachim, Russell). In Hintikka, Puhl 1994, 537–543; repr. in Woleński (1999), 233–236.
Woleński, J. (1994a). Contributions to the history of the classical truth-definition. In Prawitz, Skyrms, Westerståhl (1995), 481–495; repr. in J. Woleński (1999), 139–149.
Woleński, J. (1995). Logic and falsity. *Logica '94*, 95–105.
Woleński, J. (1995a). On Tarski's Background. In Hintikka (1995), 331–341; repr. in Woleński (1999), 126–133.
Woleński, J. (1998). Theories of truth in Austrian philosophy. *Reports on Philosophy*, 18, 13–49; repr. in Woleński (1999), 150–175.
Woleński, J. (1999). *Essays in the History of Logic and Logical Philosophy*. Kraków: Jagiellonian University Press.
Woleński, J. (1999a). Semantic conception of truth as a philosophical theory. In Peregrin 1999, 51–66.
Woleński, J. (1999b). The Semantic revolution – Rudolf Carnap, Kurt Gödel, Alfred Tarski. In Woleński, Köhler (1998), 1–15.
Woleński, J. (2001). In defense of the semantic definition of truth. *Synthese*, 126 (1–2), 67–90.
Woleński, J. (2003). Truth and satisfaction by the empty sequence. In Rojszczak, Cachro Kurczewski (2003) 267–276; repr. in Woleński(2011), 17–24.
Woleński, J. (2004). Logic, semantics and realism. In Benmakhlouf (2004), 135–148; repr. in Woleński (2011), 51–60.
Woleński, J. (2004a). First-rrder Logic: (philosophical) pro and contra. In Hendricks, Neuhaus, Pedersen, Scheffer, Wansing (2004), 369–399; repr. in Woleński (2011, 61–80.
Woleński, J. (2004b). On some formal properties of truth. In Kolak, Symons (2004), 195–207; repr. in Woleński 2011, 91–100.
Woleński, J. (2004c). Gödel, Tarski and truth. *Revue Internationale de Philosophie*, 59, 459–490; repr. Woleński 2011, 101–123.
Woleński, J. (2004d), Is truth an ontological Concept? In Marsonet, Benzi (2004), 133–159.
Woleński, J. (2004e). What are truth-bearers? *Reports on Philosophy*, 22, 34–47; repr. in Woleński 2011, 145–154.
Woleński, J. (2004f). *Aletheia* in Greek thought until Aristotle. In Adamowicz, Artemov, Niwiński, Orłowska, Romanowska, Woleński (2004a), 339–360.
Woleński, J. (2004g). Analytic vs. synthetic and a Ppiori vs. a posteriori. In Niiniluoto, Sintonen, Woleński (2004), 781–839.
Woleński, J. (2005). *Epistemologia. Poznanie, prawda, wiedza, realizm* (Epistemology. Cognition, Truth, Knowledge and Realism). Warszawa: Wydawnictwo Naukowe PWN.
Woleński, J. (2006). A generalization of Hume's thesis. *Logica Yearbook 2005*, 215–223; repr. in Woleński 2011, 155–162.
Woleński, J. (2006). Tarskian and Post–Tarskian truth. In Auxier, Hahn (2006), 647–672; repr. in Woleński (2011), 163–182.
Woleński, J. (2007). Applications of squares of opposition and their generalizations in philosophical analysis. *Logica Universalis*, 2(1), 13–29; repr. in Woleński (2011), 255–269.
Woleński, J. (2007a). Two critical contributions to the problem of truth and meaning. *Polish Journal of Philosophy*, 1(2007),137–141; repr. in Woleński (2011), 183–187.
Woleński, J. (2008). The character of T-equivalences. In Dégremont, Keiff, Rückert (2008), 511–521; repr. in Woleński (2011), 231–241.

Woleński, J. (2009). The rise and development of logical semantics in Poland. In Lapointe, Woleński, Marion, Miśkiewicz (2009), 43–59.

Woleński, J. (2010). Truth and consistency. *Axiomathes*, 20(2), 347–355; repr. in Woleński (2011), 303–311.

Woleński, J (2011). *Essays on Logic and Its Applications in Philosophy*. Frankfurt a. M.: Peter Lang.

Woleński, J. (2012). Do we need to reform the old T-scheme. *Discusiones Filozóficas*, (20)(13), 73–85; repr. in Woleński (2018), 45–50.

Woleński, J. (2013). *Ens et Verum Concertuntur* (Are being and truth convertible. A contermporary perspective)? In Dumitriu, Sandu (2013), 75–83; repr. in Woleński (2018), 107–114.

Woleński, J. (2014). Truth-makers and Convention T. In Reboul (2014), 79–84.; repr. in Woleński (2018), 155–160.

Woleński, J. (2014a). An abstract approach to bivalence *Logic and Logical Philosophy*, 23(1), 3–14; repr. in Woleński (2018), 115–124.

Woleński, J. (2015). An analysis of logical determinism. In Lapointe (2015), 423–442; repr. in Woleński (2018), 161–176.

Woleński, J. (2015a). Truth-theories in the Lvov–Warsaw School. In Brożek, Chybińska, Jadacki, Woleński (2015), 73–92.

Woleński, J. (2016). The paradox of analyticity and related issues. In Abeles, Fuller (2016), 135–138.; repr. in Woleński (2018), 205–208.

Woleński, J. (2016a). Normativity of logic. In Stelmach, Brożek, Kwiatek (2016), 169–195; repr. in Woleński (2018), 177–194.

Woleński, J. (2016b). Formal and informal aspects of the semantic theory of truth. In Kuźniar, Odrowąż–Sypniewska (2016), 55–66; repr. in Woleński (2018), 195–204.

Woleński, J. (2017). Semantic definition of truth, empirical theories and scientific realism. In Agazzi (2017), 279–294.

Woleński, J. (2017a). The story of the German translation of Tarski's book of truth. In Bar-Am, Gattei (2017), 81–90.

Woleński, J. (2018). *Logic and Its Philosophy*. Frankfurt a. M.: Peter Lang.

Woleński, J. (2018a). The semantics controversy at the Paris Congress 1935. *Philosophia Scientiæ*, 22(3), 2018, 3–15.

Woleński, J. (2018c). Some philosophical aspects of semantic theory of truth. In Garrido, Wybraniec–Skardowska (2018), 371–389.

Woleński, J. (Ed.) (1990). *Kotarbiński: Logic, Semantics and Ontology*. Dordrecht: Kluwer.

Woleński, J., Köhler, E. (Eds.) (1999). *Alfred Tarski and the Vienna Circle. Austro-Polish Connections in Logical Empiricism*. Dordrecht: Kluwer.

Woleński, J., Simons, P. (1989). *De Veritate*: Austro-Polish contributions to the theory of truth from Brentano to Tarski. In Szaniawski (1989), 391–442.

Wolf, R. S. (2005) *A Tour through Mathematical Logic*. Washington: The Mathematical Association of America.

Woods, J. (2018). *Truth in Fiction. Rethinking Its Logic*. Springer: Berlin.

Wójcicki, R. (1973). Dual counterparts of consequence operations. *Bulletin of the Section of Logic of the Institute of Philosophyf and Sociology of the Polish Academy of Sciences*, 2(1), 54–57.

Wójcicki, R. (1970). *Topics in the Formal Methodology of Empirical Science*. Dordrecht: Reidel.

Wójcicki, R. (1988). *Theory of Logical Calculi. Basic Theory of Consequence Operations.* Dordrecht: Kluwer.

Wójcik, A. S. (2010). *"Prawda" w Biblii Hebrajskiej. Analiza logiczno-lingwistyczna* (Truth in Hebrew Bible. An Logico-Linguistic Analysis). Kraków: Austeria.

Yablo, S. (1993). Paradox without self-Reference. *Analysis*, 33. 251–252.

Yaqūb, A. M. (1993). *The Liar Speaks the Truth. A Defence of the Revision Theory of Truth.* Oxford: Oxford University Press.

Zinoviev, A. A. (1963). *Philosophical Problems of Many-Valued Logic.* Dordrecht: Reidel.

Author Index

J. Woleński, *Semantics and Truth*, Logic, Epistemology, and the Unity of Science 45, https://doi.org/10.1007/978-3-030-24536-8

Subject Index

J. Woleński, *Semantics and Truth*, Logic, Epistemology, and the Unity of Science 45, https://doi.org/10.1007/978-3-030-24536-8