Volume 3, no. 1 (October 1992)

GÖDEL'S FIRST WORKS, 1929-1936: MATHEMATICS WITHOUT PHILOSOPHY?

Review of Kurt Gödel, *Collected Works. Volume 1. Publications 1929-1936*. Editor-in-chief Solomon Feferman. New York and Oxford: Oxford University Press, 1986. xvi + 474.

Francisco A. RODRÍGUEZ-CONSUEGRA

Departamento de Lógica, Historia y Filosofía de la Ciencia
Facultad de Filosofía, Universidad de Barcelona
E-08028 Barcelona, Spain

I

This is a very special book. It contains the official and definitive German and English versions of one of the most famous and influential papers in the entire history of logic, written by one of the great geniuses in the intellectual history of the human race. Besides, it contains all the remaining publications by Gödel in the period covered, in both languages (when necessary). It is the first volume of a long series which intends to critically edit all of Gödel's publications, together with the most significant unpublished writings. In addition, the team of six editors is impressive: the three mathematical logicians Solomon Feferman, Stephen Kleene and Robert M. Solovay, and the three historians of logic and mathematics John W. Dawson, Jr., Gregory H. Moore and the late Jean van Heijenoort, all of them working under the auspices of the Association for Symbolic Logic. They have been in charge of the completion of Gödel's publications, the general introduction, the chronology, most of the introductory notes, the textual notes, and the selection of the references (not to mention other minor sections which will be described below). Moreover, there are other important names in charge of particular introductory notes, including B. Dreben, A.S. Troelstra, W.D. Goldfarb, W.V. Quine, J. Webb, R.L. Vaught and R. Parikh. We have then before us a very important book, and to write a critical study of it is a complex task.

My strategy is the following. First, I will describe the content of the book from the point of view of Gödel's works, making some comments on the way they have been organized for publication (II). Then I will describe the different sections other than the introductory notes, adding more comments about their usefulness (III). Finally I will concentrate on the introductory notes, first to indicate their general features (IV), and secondly to describe and discuss the content of some of them, precisely those devoted to the philosophically most significant of Gödel's writings included in this volume (V-IX). At this point, I must state what will be my main criticism of this edition, which I have tried to capture in the title I have chosen: it seems to me that the philosophical motivations and implications of Gödel's results are not elaborated on enough.

This conviction depends upon a general thesis which I think, if true, is very important: that Gödel was mainly a philosopher searching for mathematical results to illustrate, if not to prove, the truth of a philosophical position, apart from the mathematical interest of those results. That philosophical position was a twofold kind of Platonism, i.e. ontologically, the belief in the existence of separate and transcendent abstract entities, and epistemologically, the belief in the existence of a human intuition allowing us somehow a direct access to those entities. Also, I would suggest that in spite of this particular criticism, this book is going to be the official, canonical and of course indispensable source for the study of Gödel's works for many years. This is appropriate, because the work is very competently done and contains the essential instruments which the Gödel scholar, or any other person interested in logic, mathematics or, more partially, philosophy, could possibly need.

II

The writings by Gödel which are published here can be divided into three categories: (i) major writings; (ii) minor writings; (iii) reviews. The major writings are the following. First, the article from 1930 on completeness of first order logic, together with the original dissertation from 1929, which is extremely interesting, not only because of the several modifications which appear only in the article, but also because the dissertation contained a fascinating philosophical and – mildly – polemical introduction which was later deleted from the actual publication. This is the first time this introduction is published, so the reader interested in historico-philosophical matters will doubtlessly study it carefully. Secondly, we have the star of the volume: the bilingual version of "On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems I" (1931). This has been printed many times, but it is nice to have it at last together with the rest of its natural neighbours, like the several summaries and notes which Gödel published on several occasions, and, remarkably, Gödel's contribution to the Königsberg discussion of 1931 containing the first announcement of the incompleteness results. Thirdly, there are the Princeton lectures from 1934 (first published in 1965) containing some developments of the incompleteness results, in particular a section entirely devoted to some relationships between those results, the paradoxes and the theme of truth, which was unfortunately missing from the original paper of 1931. Lastly, we have the paper from 1933 in which Gödel proved that – in a special sense – intuitionistic arithmetic is not really narrower than classical arithmetic (see above). As we shall see below, every one of these writings can be interpreted as having important philosophical motivations and implications which are not always adequately treated in the corresponding introductory notes.

The minor writings are very different in nature, extension and importance. Among them, we can find a few which are closely related to the ones I have mentioned above, that is, closely related to the themes of completeness, incompleteness and intuitionistic logic and arithmetic. Also, there are two from 1932 and 1933 devoted to the decision problem, as well as one from 1936 on the theme of the length of proofs. Incidentally, it has to be said for the decision problem that, although Gödel proved that certain classes of formulas of first order logic are decidable, his extension of the same argument to the predicate calculus with identity was shown to be false in 1984 by Goldfarb, who, fortunately, is in charge of the corresponding introductory notes. This is the only technical mistake Gödel made in print, as far as we know. If we add that some years later he withdrew at least one paper from publication when other mistakes were found (see p. 27), we can arrive at the comforting conclusion for the ordinary logician and philosopher that he was, after all, human too. Also, there is a series of three short notes from 1932 and 1933 on several themes of the propositional calculus, one of them on the possibility of precisely defining the expression "*p* analytically implies *q*", which seems to me to show Gödel's – presumably philosophical – interest in analyticity. Finally, we have five short notes on geometrical themes from 1933, which, although hardly interesting for the usual Gödel scholar, can be seen as a further sign of his impressive intellectual capacities.

The third and last category is constituted by the reviews. We have here a long series (more than thirty) of short reviews of books and articles, some of them very interesting in principle because of the obvious connection of the authors reviewed to Gödel's results, e.g. papers by Heyting, von Neumann, Skolem, Quine, Carnap, Church, Hahn and Hilbert. The reviews are however rather disappointing for the ordinary reader, who may hope to find some of Gödel's *opinions* about some of the fundamental ideas of those authors. For instance, as soon as I saw the book for the first time I immediately read the reviews devoted to Carnap and Hilbert, with the hope of reading something really fascinating, and even devastating, on Carnap's logical syntax of language (as he tried to present a conventionalist alternative to any Platonistic account of mathematics) and Hilbert's formalist programme (as he tried to provide complete formalizations of mathematics, and perhaps dispense with some of the traditional abstract entities needed by the Platonist). Unfortunately, Gödel completely avoided any opinions and limited himself to a short description, as objective as possible, of the content of the several papers. This however is interesting, for together with the fact that he deleted the introduction to his dissertation from the article he wrote for publication, it provides a clear sign of one of the more remarkable features of Gödel's behaviour: his extreme caution in expressing philosophical views. This was probably due to his highly idiosyncratic personality, which made him avoid philosophical polemic, particularly in an era when nominalism and empiricism became what he once described as "the prejudices of the time."

As a whole, most of these materials can be regarded as practically new, as many of the notes and reviews were unknown and have been rediscovered only very recently by John Dawson, who not only organized the Gödel *Nachlaß* in 1982-84 and wrote interesting papers on Gödel and the reception of his results, but also is now writing what I am sure is going to be an impressive biography of Gödel. One completely new item is the first printing of Gödel's original dissertation. Now we have for the first time a splendid, and probably complete, anthology of the period covered.

Let me call attention to a couple of critical points. All bibliographical references are made by the convenient, and highly historically informative style of author/year of publication. This is fine for books and papers by authors other than Gödel himself. But in the case of Gödel's own writings, a glance at the contents of the book may mislead the reader into seeing the several references as being somehow equivalent to each other, while actually there really are major works intermingled with short notes. It would have been more informative to adopt another system at least for major works, leaving the year/letter style for minor papers, notes and reviews. As for the overall decision to adopt the year of publication/ presentation, although correct, some exceptions might perhaps have been allowed. For instance in the case of *1931* (the paper containing the incompleteness theorems which appeared in January 1931) and *1931a* (the contribution to the Königsberg symposium, which took place in September 1930), the reader is led to believe that the second is later than the first, and this is true if we mean that the second was *published* in 1931, after the first had already appeared. But it is also true that in *1931a* Gödel implicitly referred to, and actually announced his results of *1931*; so in this case, choosing the date of the *presentation* would have avoided certain problems.

III

I come now to sections other than the Gödelian materials themselves and the introductory notes. A useful and friendly preface is provided by Professor Feferman in which he tells us the history of the project. There we learn, among many other things, that the encouragement of Gregory H. Moore and Jean van Heijenoort was of "pivotal importance" for its development. A useful section "Information for the Reader" follows, providing some comments explaining the criteria which have been chosen for organizing the introductory notes, references, textual notes, translations, etc. (I shall say something about these criteria in this and the next section). Then, after the copyright permissions, a note of gratefulness to the late Julia Robinson, and the "Contents," we arrive at the first major section of the book, Professor Feferman's thirty-six page essay, "Gödel's Life and Work."

In short, I think this essay is the best short account of the global Gödel available in print, so it can be used as a splendid introduction to every aspect of our theme. The first part, devoted to Gödel's life and career, is preceded by the legendary words of Adele Gödel (Gödel's wife, who never studied philosophy, let alone mathematical logic) after attending Gödel's highly technical Gibbs Lecture: "Kurtele, if I compare your lecture with the others, there is no comparison." The short biography is very accurate and, at the same time, very interesting for the reader. Also, Feferman's obviously sympathetic attitude does not led him to hagiography. He does not avoid discussing the more obscure aspects of Gödel's personality, for instance his more or less frequent nervous depressions in the thirties, his hypochondriacical tendencies and his fascination with demonology.

One particularly significant feature of Gödel's life is also clearly noted: the fact that he devoted himself almost entirely to philosophy from 1943 onwards (when he was only 37 years old!), that is, after his main results in metamathematics and set theory were achieved, and after he failed to prove the independence of the axiom of choice and the generalized continuum hypothesis from the standard axioms of set theory (which was proved by Cohen in 1963). Usually, we would describe as a philosopher a person who devotes 13 years to a particular activity and the rest of his life, 25 more years, to philosophy. The fact that Gödel is not so regarded is because he published highly influential results only in mathematical logic, and published very little on his philosophical "results." This is one of my arguments supporting the thesis that Gödel was mainly a philosopher, who devoted a few years to a certain branch of a particular exact science, mainly because he expected to find there exact results to be presented as "proofs" that his philosophical doctrines were true.

The second part of Feferman's essay, devoted to Gödel's works, thought and influence, is divided into four sections. In the first one, all the materials of the first two volumes are grouped, very judiciously, into thirteen systematic categories, each one of them provided with excellent summaries of their contents. The second section groups the materials which are extant in the *Nachlaß* into four categories: unpublished manuscripts, individual lecture texts, lecture notes, and notebooks. Unfortunately, there are no summaries here, as the reader is directed to "a succeeding volume," while no detailed prospective of that, or other succeeding volumes, is offered. Section 3 is gives a fine and clear exposition of Gödel's philosophy of mathematics. However, Feferman's treatment of Gödel's doctrines consists of seeing them as "emerging" from the publications (the metaphor is used twice: pp. 22 and 30), which will probably be interpreted by most readers as if Gödel's philosophy of mathematics were some kind of "secondary" product implicit there. Feferman is clearer in section 4, where he cites Gödel's realist position as a strong motivation and an important part of the reason for his success (p. 32). I think that Gödel's philosophy of mathematics was only part of a consistent whole, and that it can be interpreted not only as a fortunate motivation, but also as the main objective of his intellectual life. In short, and very roughly: the belief in transcendent entities is not only good because it leads us in the right direction to find metamathematical results, but these results are somehow the proof that those entities exist.

There are only two more sections remaining: the useful Gödel chronology by Dawson, immediately before the first of Gödel's papers, and the accurate "Textual Notes" (containing the necessary references to the copy-texts, the paginations and the replacements which the editors have introduced in the original texts), immediately after the last paper.

The bibliographical references and the index are doubtlessly very important in evaluating the usefulness of a work like this. The list of references (about whose author or authors we are told nothing) has been apparently constructed according to a threefold set of criteria (p. vi): Gödel's bibliography, and items referred to by Gödel or by the editors or other authors of the introductory notes. So we would expect that every item in the list, if not by Gödel himself, was referred to by some member of the team somewhere in the book. But this is by no means the case. Thus, the index contains many items as appearing only in the list of references. Russell's works are a good example: while the corresponding entry (p. 447) offers us ten titles, we learn in the Index (p. 471) that only three of them are actually referred to in the body of the book. Since readers do not have at their disposal any criterion to determine why certain titles appear and why others do not, they can conclude that a certain degree of arbitrariness is present.

More concretely, there are a few significant omissions, which I have chosen more or less at random (there are *many* more). First, it was the Spanish edition of Gödel's [*1981*] *Obras Completas*, edited by Jesús Mosterín, which was the first attempt in any language to collect Gödel's publications. Second, is what I think can be described as the most famous first introduction to Gödel's incompleteness theorems: Nagel and Newman's [*1958*] *Gödel's Proof*. (I recently told Professor Dawson about these two omissions. He told me that the first one was not intended, while the second was the result of an explicit decision, as the famous booklet is simply a poor introduction.) Third, the very influential article, [Lucas *1961*], which gave rise to a whole series of replies, and is now indispensable in evaluating some of the philosophical implications of Gödel's incompleteness results for the problem of the nature of the human mind. Fourth, the massive [Hofstadter *1979*]. I think that these last two omissions are further consequences of the rather unfair treatment of philosophical matters which is present in this work (I will return to that below).

I have only one criticism to make of the index: it is an index only of names. It is perfectly legitimate and understandable, but the value of the book for the scholar would have been much greater had she/he been able to follow the track of certain key terms throughout the very dense xvi + 474 pages.

IV

The introductory notes to Gödel's papers and reviews try to provide (p. vi): (i) the historical contexts; (ii) explanations of the contents; (iii) discussions of further developments; (iv) critical analyses (admittedly, only in some cases). The Editors say that no uniformity has been imposed, although every note has been accepted by them as a whole. However, the notes meet these common purposes in varying degrees. Since the authors are highly competent in their field the results are usually extremely good in spite of the lack of uniformity. However, as for the general degree of difficulty of the notes, I think that very few readers will be able to understand them without, at the same time, being able to understand the papers themselves. So from the point of view of the "explanation of the contents," the notes seem to me to have been composed rather for the specialist than for the general reader interested in Gödel coming from other fields of science or philosophy. It is in this sense that the four purposes have been planned more around the interests of mathematical logicians than of philosophical logicians or simply philosophers. Thus, consideration of the historical *philosophical* context, the *philosophical* implications of the contents, further *philosophical* developments, and critical *philosophical* analyses, are unfortunately almost absent from the introductory notes, apart from some allusions, or the use of a few standard philosophical terms and labels.

From the point of view of the reader, it is very easy to know at first glance whether one is looking at an introductory note or one of Gödel's writings: in the first case there is always a vertical line on the outside which runs the full length of the page. Also, every note begins with the common form "Introduction to.... (the key/s to the paper/s)," and we are told that every one of Gödel's writings has the corresponding introductory note. However, some readers will perhaps be puzzled in trying to find the note to a particular paper or review, as in some cases the editors have grouped, under a particular introductory note, the material corresponding to several writings. This is not a serious problem; it is only a matter of some inconvenience. But the inconvenience increases as: (i) the papers and reviews are printed chronologically, so the materials covered by the same introductory note may be located in very different places in the book; (ii) Gödel's intellectual production in these years consisted mainly of many short publications a year (e.g., from 1932 we have *1932*, *1932a*, ...*1930o*, that is 16 items!), so sometimes it is difficult to find quickly what is needed.

The Editors must have considered the case where the material covered by the note is not located immediately after it. We may read, for instance, a short "⟦The introductory note to... can be found on page..., immediately preceding...⟧" before the first line of the paper in question. However, the treatment is not always uniform, as there are instances where common introductory notes simply say almost nothing about some of the writings. One example of that is [Gödel *1932b*], "On Completeness and Consistency" (p. 235 ff), which is only mentioned in the "Introductory Note to *1930b*, *1931* and *1932b*" (p. 126 ff), by Professor Kleene, who devotes only two and a half lines to it (p. 126). The reader especially interested (this was my case), is forced to go to the index or the bibliographical reference to learn something else about that particular paper, and what can be found is only the place and time of publication. Perhaps this is all the information relevant to this paper, but I cannot help being convinced that a much better system would have been: "a paper, an introductory note."

V

From here on, I will concentrate on five of the introductory notes devoted to the five papers or notes which are philosophically more important, without forgetting that the importance of an idea in history of philosophy is measured by the influence exerted by its direct or indirect influence on the intellectual period considered.

I will begin with the introductory note to Gödel's dissertation of 1929 and related materials (p. 44 ff), by Burton Dreben and Jean van Heijenoort. Let me first say clearly how much I admire this note, which can be described as historically superb. Its first section is devoted to the historical background of the idea of completeness, that is: the Frege-Russell inability to imagine a non-universalist and comprehensive logic; the absence of the notion of formal system in Peirce-Schröder; the work by Post and Bernays demonstrating the semantic completeness of the propositional calculus; the first appearance of the same question for quantification theory in Hilbert-Ackermann.

Missing here is only some philosophical explanation for the reason why Frege-Russell thought it impossible to look at logic from outside it. I think the crux of this question was their belief that any attempt to do so leads to unsolvable paradoxes belonging to the same family as Bradley's paradox against relations. Thus, to escape from logic presupposes defining its main notion, that of logical form, which is a kind of relation, in such a way that logical form appears to be a genuine concept (a "term") related to other concepts. But as soon as we say that, we are forced to explain the relation between forms and non-forms, which leads us to other forms, that is, to an infinite regress. As we shall see below, a parallel philosophical problem appears in relationship to Gödel's incompleteness results and the question of the expressibility of truth and syntax in the same language.

The study of the introduction to the dissertation is philosophically the most interesting section. The authors accurately explain Gödel's main argument against the formalist belief that the consistency of an axiom system implies the existence of the corresponding mathematical concept: "it assumes that no formula in the notation of the system can be proved to be undecidable" (p. 49), that is, it assumes that every problem is solvable. Then, after criticizing Gödel's example of an axiom system for the reals with two non-isomorphic models for not mentioning the distinction between first and second order logic (which presumably means that the argument is not correct, although this seems to be only suggested; see [Feferman *1984*, 551, footnote 7], who describes the argument as "inconclusive"), Dreben and van Heijenoort point out that Gödel's criticism of the formalist belief goes further than Brouwer's, for it is not only that we cannot simply assume the solvability of every mathematical problem, but also "that we may be able to prove the unsolvability of some mathematical problem" (p. 50), with which the incompleteness results seem almost to be announced.

However, the authors do not mention the fact that it was precisely Gödel's completeness theorem which, in demonstrating that if an axiom system is consistent then it has a model, proved that consistency is somehow equivalent to existence. So his reluctance to admit this equivalence can be interpreted as meaning simply that we cannot assume it *before* having obtained the completeness result (on that point see [Feferman *1984*, 551 f]). Thus, the following step towards proving incompleteness can be seen as pointing out the definite limits of consistency from the ontological point of view: before incompleteness was proved, the equation of consistency with existence could be still maintained. Perhaps this may help us to understand Gödel's final decision to delete the introduction from the printed version, a problem which the authors leave open (as Feferman did in his [*1984*, 552]).

The main section of this introductory note is the technical one devoted to the proof of completeness itself, whose precise progress is completed by comparisons with some of Skolem's former ideas (where important letters by Gödel are quoted), as well as with Herbrand. A significant summary of these comparisons is the following: "let us consider the sequence: Frege-type formal system, proof procedure, set-theoretic validity. Skolem connects (at least implicitly) the second and third terms of the sequence, Herbrand the first and second, Gödel the first and third" (p. 55). The section finishes by making some comments on further developments by Gödel himself and Henkin, and so does the last short section. As for the *introductory* nature of this note: I would have liked to read somewhere a few simple paragraphs emphasizing the intuitive gist of the proof and the main standard consequence for logical truth and provability, very roughly like the following, which can be found in standard introductions to logic.

As for the basic idea of the proof, we can say that the completeness of quantification theory (that every one of its valid formulas is a theorem) is proved by showing that every valid first-order formula can be reduced to a valid propositional formula from which it can be derived. Then, as the propositional calculus had already been proved to be complete (by Bernays), every valid propositional formula is also provable, so every first-order formula (which can be derived from one of the former) is provable too, and first order quantification theory is complete. In short, what Gödel proved was that first-order logic is complete if propositional logic is complete.

As for the standard interpretation of the completeness result, we can say that it shows that for first order languages a formula is logically true iff it is logically provable, i. e. truth is equivalent to provability, which means that the corresponding proof procedures are complete. This can be seen as philosophically relevant, especially if we express the same idea in other – and later – terminologies (and always for first-order logic): logical consequence is equivalent to derivability; semantic consistency is equivalent to formal consistency; every semantically valid formula (true in every system) is syntactically demonstrable (derivable with no premiss); in short: semantics is equivalent to syntax.

VI

Gödel's contribution to the Königsberg conference of 1930, published in 1931 after his incompleteness results, was the next natural step after his former unpublished criticisms against formalism that we have discussed above. John Dawson's introductory note to that contribution comes from his excellent paper on the same topic [Dawson *1984*], in which he provided a wider commentary, as well as the English translation of the contributions by other participants (Hahn, Carnap, von Neumann).

The most interesting point of Gödel's contribution, apart from containing the official announcement of his completeness result, is the continuation of his dissertation criticisms against consistency of an axiom system as a criterion for existence of a mathematical concept, which is now based upon an implicit announcement of the incompleteness results. However, the idea that is expressed here is rather against consistency as a criterion for the truth of every theorem provable in the system. Dawson summarizes the argument by saying that "contrary to Carnap, Gödel argues against adopting consistency as a criterion of adequacy for formal theories," for one could perceive, finitarily and contentually, "that a statement provable within some consistent formal system is nonetheless false," as "one can give examples of contentually true propositions that are unprovable within the formalized framework of classical mathematics" (p. 197). This summary, being essentially correct, leaves unexplained several interesting points, doubtless because for Professor Dawson they are too obvious. But I think for the general reader these points should have had more elaboration.

First, the allusion to Carnap. I think it is correct, providing Carnap's philosophy of mathematics at that time can somehow be identified with that of Hilbert and his followers, against whom Gödel's remark is explicitly directed (he speaks of "the formalist view," p. 201). For Hilbert, a proof of consistency for a system would guarantee the truth of every theorem provable within it. (In this he directly contradicted Frege, for whom the true situation was just the reverse: a set of true axioms is already trivially consistent, so there is no need of any proof of consistency.) If we can prove that there can be a consistent system from which we can derive a false theorem, then the proof of its consistency is not enough. And this is what Gödel says can be proved through his argument.

Secondly, it seems to me that the link between the fact that a false statement can be provable within a consistent system, i.e. that of classical mathematics (which we take to be so), and the fact that contentually true propositions can be unprovable within this system is not emphasized enough. I think this link is established by the conclusion to Gödel's contribution: if there is some contentually true proposition which is unprovable in classical mathematics, then "if one adjoins the negation of such a proposition to the axioms of classical mathematics, one obtains a consistent system in which a contentually false propositions is provable" (p. 203). Therefore, we can add, consistency does not guarantee truth, as consistent systems can contain false statements.

In this context it may be worth recalling Russell's early strong opposition to consistency as an ultimate criterion for truth which appeared in several places in his work, for instance, against Bradley's coherence theory of truth, and against "implicit definitions" as being able to guarantee the existence of the concept supposedly defined by them. Contrarily, for Russell (i) consistency is not enough, as there can be consistent systems constituted by false statements; (ii) we need first prove the existence of the concept involved.

One of the philosophical morals of Gödel's global argument can be stated this way: if consistency does not guarantee truth or existence, we cannot create objective concepts simply by constructing axiom systems, i.e., concepts exist or not on their own. A second moral could be: any attempt to interpret mathematics in a formalistic – then somehow conventionalist – way is hopeless, precisely because we by no means can forget mathematical intuition, nor replace it by a proof of consistency which can be understood in an algorithmic, mechanical way. Needless to say, this is one of the main lines of the philosophical implications of the incompleteness results. So there should be some reference to it in this introductory note, as it was the first place in which Gödel made public some of those results, and he did it in a full philosophical context.

VII

I come now to the kernel of Gödel's work in this period: his celebrated paper of 1931 containing the incompleteness results for which Professor Kleene wrote the exquisite introductory note. It would be difficult to find a person better qualified for this task.

There is a clear summary of Hilbert's programme (a programme based on the hope that consistency and completeness are attainable for formalized mathematics), then a short account of Gödel's incompleteness results showing the impossibility of meeting both requirements at the same time. This is followed by a brief exposition of the central ideas of Gödel's first theorem. Next the exposition of Gödel's paper begins, which is structured according to its four original sections, and constitutes the bulk of the introductory note. Finally,three pages are devoted to further developments in the field, particularly several attempts to provide finitary consistency proofs.

As a whole I think the note is technically perfect, although it is very difficult to comprehend if we regard it as a first "introduction" to the theme. From an historical viewpoint I think the situation is unfortunately different. It is true that Professor Kleene makes useful references to Gödel's development, starting from the discovery of the undefinability of arithmetical truth in arithmetic (p. 127), as well as Gödel's first treatment of the question at the Königsberg meeting (pp. 135, 137). Also, he devotes a footnote (p. 137) to refer the reader to the literature on the reception of Gödel's results. However, there is no discussion of the possible problem of priority for what now is known as "Tarski's theorem" of the undefinability of truth in the same language, which is particularly needed as Kleene seems to accept that Gödel already had the idea in 1930. This is needed especially since Kleene himself devotes some interesting remarks to interpret Gödel's rather cryptic footnote 48a, by writing (p. 135): "Implicit in this remark is that the adjunction of higher types to a formal system permits one to define the notion of truth for that system, then to show that all its provable sentences are true, and hence to decide the sentence shown in Theorem VI to be undecidable in the system" (see below for further discussion of this point).

Moreover, as for Gödel's development of the incompleteness idea, we do not find any treatment of the historico-philosophical information available in the literature at the time when the book appeared (mainly [Feferman *1984*] and [Dawson *1984*; *1985*]). For instance, there is no allusion to Finsler, the only precedent to Gödel. It is true that he was mistaken, especially as he thought absolute formal undecidability could be proved, but I think a few lines about him would have been useful for the reader to have another instance of the historical fact that original ideas always have some sort of precedent. As for the reception of Gödel's results, I am convinced that a simple footnote to the relevant literature is not enough. Some treatment of the main problems forestalling full acceptance is also necessary.

From the philosophical viewpoint things are, in my opinion, not much better. I do not have enough space to devote even a few lines to every missing point I think would have been relevant here, so I shall only mention what I think are the most important of them:

– the actual historical reception of Gödel's results from the view-point of at least a few well-known names: Hilbert, Brouwer (see [Wang 1987], Russell (see [Rodríguez-Consuegra *1992*]), Wittgenstein (see Shanker *1988*, 155-256]);

– Gödel's conviction of the objectivity of mathematical truth, as opposed to mere provability;

– implications for logicism: impossibility of a unique global account of formalized mathematics;

– possibilities of partial reconstructions of the logicist programme by assuming Gödel's results;

– implications for intuitionism: Gödel's interest in giving a proof acceptable to the intuitionists, as he had available earlier a general proof of incompleteness which was not acceptable (letter to Zermelo of 1931; see [Grattan-Guinness *1979*]);

– implications for Platonism: are Gödel's results an argument to vindicate Platonism? (see [Myhill *1952*]);

– implications for realism: are Gödel's results an argument to vindicate philosophical realism? (see [Dummett *1963*]);

– implications for conventionalism: can Carnap's syntactical program of 1934 be refuted by Gödel's results, though it was explicitly constructed by starting from them? (see [Bohnert *1975*]);

– implications for the analytic-synthetic question: could Gödel's results have helped to give rise to Quine's famous attack against the analytic/synthetic distinction? In particular, can those results be interpreted as having proved that some synthetic knowledge is possible, or at least that mathematics cannot be analytic? (see [Copi *1949*]; [Körner *1967*]);

– are Gödel's results relevant for the problem of the nature of the human mind, especially concerning algorithmic possibilities? (see above, section IV, and below, section IX).

I am fully aware of the limitations imposed on an introductory note, and of the fact that often mathematical logicians are not very fond of philosophical questions. In particular, I know that Professor Kleene has written else-where [*1976/78*, 77] that he does not consider himself qualified to give an account of Gödel's work in philosophy. But I cannot help feeling that since this is the official edition of Gödel's works, the introductory note to such an incredibly rich and influential paper should have been explicitly treated some of its most important philosophical implications, regardless of the view that Gödel was considered to be only a mathematical logician.

VIII

Gödel's interest in intuitionism hardly needs to be emphasized As I pointed out before, one of the main motivations for his particular way of proving formal undecidability was that the proof would be accepted by intuitionists. In the period covered by this book, Gödel wrote several short papers which have to do with intuitionism. I think two of them are relevant to understanding Gödel's philosophical motivations, because they provide further evidence that most of the time he may have regarded mathematical results also as the basis for his philosophical convictions. The papers were written in 1933 and are entitled "An Interpretation of the Intuitionistic Propositional Calculus" and "On Intuitionistic Arithmetic and Number Theory."

For the first paper, Professor Troelstra writes in the corresponding introductory note that it contains an interpretation of intuitionistic propositional logic in an extended system of classical propositional logic. Then, after describing the details and further developments, he adds: "For Gödel, the interest of his result presumably lay in the fact that it gave for IPC an interpretation which was meaningful also from a non-intuitionistic point of view" (p. 299). Regarding the second paper, Gödel first showed that classical propositional calculus is a subsystem of the intuitionistic one, and then that something similar holds for arithmetic and number theory. On the implications of this thesis, Professor Troelstra, who is also the author of the corresponding introductory note, only writes: "Gödel concludes the note with the observation that his results show that intuitionistic arithmetic is only apparently weaker than classical arithmetic" (p. 283).

But, we can ask, why was Gödel so interested in this sort of reducibility of classical to intuitionistic formal systems? To provide some answer to this question is relevant for intuitionism as well as for formalism and logicism (which were more inclined to use classical logic and arithmetic) and are mainly *philosophies of mathematics*. A part of the answer is provided by Gödel's own remark at the end of the last paper, indicating that "an intuitionistic consistency proof for classical arithmetic and number theory" is involved in this result (p. 295). This is true, we can add, because the reduction proves any theorem of classical arithmetic to be *classically* equivalent (that is why intuitionists can reject the result!) to a theorem of intuitionistic arithmetic, so we can no longer think that within intuitionistic arithmetic we can feel more sure of avoiding contradictions than within classical arithmetic. This is perhaps too obvious, but once it has been explicitly noticed, one can immediately understand that the intuitionists' extreme caution in the fully philosophical search for a secure basis for mathematics, could be overcome, at least to some degree, through a technical result. If so, this can be interpreted as another instance of Gödel's hope of providing technical arguments for philosophical polemics.

IX

To complete this critical study, let me say something about Gödel's Princeton lectures of 1934. Professor Kleene again is responsible for the excellent introductory note to those lectures, which he personally attended lectures and, together with J.B. Rosser, prepared the notes they took for publication.

Kleene explains, one by one, the technical content of the nine sections of the text, as well as the "Postcriptum" from 1964, showing that the lectures cover more territory than the original paper of 1931, especially regarding some improvements in the proof that every primitive recursive function is representable in the system, and the definition of a new and important mathematical idea, that of a general recursive function. Kleene goes on to describe further developments, where the names of Church, Turing, Rosser, Tarski, Mostowski, Robinson and Kleene himself are introduced to establish the relevant links with Church's thesis (that every effectively calculable function is general recursive); Church's undecidability theorem; Turing machines and computability; Rosser's replacement of Gödel's ω-consistency by simple consistency; Tarski, Mostowski and Robinson's work on undecidable theories; and Kleene's generalizations of Gödel's first theorem.

Three points, all related to the philosophical issues which I have referred to above need to be made. First, I find Kleene's treatment of section 7 of Gödel's text entirely insufficient. He only says that it contains a discussion of the relation of his arguments to the paradoxes, where Gödel credited Carnap and Tarski for the general self-referential lemma and the undefinability of truth in the formal system. As for crediting Tarski, I think some discussion of the possible priority is necessary in view of the fact that it may be said that Tarski's theorem of the undefinability of truth in the same language can be credited instead to Gödel. Some of the grounds for believing this to be the case are the following (for the full historical evidence see mainly [Grattan-Guinness *1979*], [Wang *1987*] and [Coffa *1987*]): (i) it was known to Gödel as early as 1930; (ii) it was proved by Gödel in 1931, at least before Carnap (and perhaps before others); (iii) it was proved in the letter to Zermelo of 1931; (iv) Tarski arrived at the discovery of the theorem only after having read about Gödel's undecidability results; (v) even so, Tarski did not include *clearly* the undefinability of arithmetic truth in arithmetic in the 1933 monograph – where Theorem I of §5 is usually, and mistakenly, cited on this point – but rather in the Postcript from 1935; (vi) Tarski's himself, in the footnotes to his famous monograph, recognized his dependence upon Gödel, and so did Gödel, in an unpublished letter to Burks of February 6,1964.

Secondly, section 7 contains something extremely valuable from the historico-philosophical viewpoint: Gödel's first proof in print (Princeton 1934) that truth cannot be expressed in the same language, which is immediately applied to arithmetic, is Gödel's first statement in print of the heuristic argument for the existence of undecidable formulas. The argument, which already appeared in his letter to Zermelo of 1931, reads as follows: "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 $\alpha \neq \beta$ and if we assume $\beta \subseteq \alpha$ (i.e., every provable formula is true) we have $\beta \subset \alpha$, i.e., 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" (p. 363). I think this great argument deserved a more complete treatment.

Gödel's use of his technique of arithmetization in reconstructing the same argument for undecidable propositions to show that a particular undecidable proposition can actually be constructed through the expressibility of the syntax of a language in the same language (1931) seems to me of the utmost philosophical importance too. Thus, he solved a very old philosophical problem from the time of Bradley and Frege, for whom only the universal logical language was inescapable. Russell tried to overcome the problem through his theory of types, but Wittgenstein rejected it through his dictum that we cannot speak about language at all, as there is really only one language; therefore Russell wrote in his introduction to the *Tractatus* that there is at least one escape: one can speak of a language by constructing a metalanguage. Then Gödel (and Tarski) showed that we can express the syntax of a language in the same language without contradiction. Thus, the only thing we cannot express in a language is its semantics, which requires a metalanguage.

Finally, Gödel's "Postcriptum" of 1964 contains the first statement in print of his full philosophical conviction: that the incompleteness results, together with the implications of the ideas involved in the concept of Turing machines, "do not establish any bounds for the powers of human reason, but rather for the potentialities of pure formalism in mathematics" (p. 370). Gödel had developed this statement in his Gibbs lecture of 1951, as well as in his contributions to [Wang *1974*]. This is a further point which, I am convinced, requires a more adequate treatment. To summarize my underlying argument, which has been applied throughout this study: if some philosophical ideas were true and important for Gödel, they have to be made known to everyone interested in any facet of his thought.

ACKNOWLEDGEMENTS. I am grateful to John Dawson for all the general information he has given me on the global project of Gödel's *Collected Works*, as well as on some of the points I mention in section III. Also, my thanks are due to Ignasi Jané, who read an earlier version of this critical study and made remarks which led me to introduce several improvements, both in content and in exposition. Finally, I am indebted to Francine Abeles, the Book Review Editor, for improving my English.

REFERENCES

BOHNERT, H. 1975. *Carnap's logicism*, J. Hintikka (ed.), *Rudolf Carnap, logical empiricist*, Dordrecht, Reidel.

COFFA, A. 1987. *Carnap, Tarski and the search for truth*, Noûs, **21**, 547-572.

COPI, I. 1949. *Modern logic and the synthetic a priori*, Journal of Philosophy **46**, 243-245.

DAWSON Jr., J. W. 1984. *Discussion on the foundation of mathematics*, History and Philosophy of Logic **5**, 111-129.

—. 1985. *The reception of Gödel's incompleteness theorems*, in P.D. Asquith and P. Kitcher (editors), *PSA 84: Proceedings of the biennial meeting of the Philosophy of Science Association* (East Lansing, Michigan, Philosophy of Science Association, 1985) vol. **2**, 253-271. Reprinted in [Shanker *1988*, 74-95] and in T. Drucker (editor), *Perspectives on the history of mathematical logic*, (Boston/Basel/Berlin, Birkhäuser, 1991), 84-100.

DUMMETT, M. 1963. *The philosophical significance of Gödel's theorem*, Ratio **5**, 140-155.

FEFERMAN, S. 1984. *Kurt Gödel: conviction and caution*, Philosophia Naturalis **21**, 546-562.

GÖDEL, K. 1981. *Obras Completas*, edited by J. Mosterín (Madrid, Alianza); second, really "complete," edition, 1987.

GRATTAN-GUINNESS, I. 1979. *In memoriam Kurt Gödel: his 1931 correspondence with Zermelo on his incompletability theorem*, Historia Mathematica **6**, 294-304.

HOFSTADTER, D. R. 1979 *Gödel, Escher, Bach: an eternal golden braid*, New York, Basic Books.

KLEENE, S. C. 1976/78. *The work of Kurt Gödel*, Journal of Symbolic Logic **41**, 761-778; *Addendum*, **43**, 613. Reprinted in [Shanker *1988*, 96- 114].

KÖRNER, S. 1967. *On the relevance of post-Gödelian mathematics to philosophy*, I. Lakatos (editor), *Problems in the philosophy of mathematics* (Amsterdam, North-Holland), 118-133.

LUCAS, J. R. 1961. *Minds, machines, and Gödel*, Philosophy **36**, 112-127.

MYHILL, J. 1952. *Some philosophical implications of mathematical logic*, Review of Metaphysics **6**, 165-198.

NAGEL, E. & NEWMAN, J. R. 1958. *Gödel's proof*, New York, New York University Press.

RODRÍGUEZ-CONSUEGRA, F. A. 1992. *Russell, Gödel and logicism*; Paper read at the Wittgenstein Symposium on Philosophy of Mathematics, Kirchberg, Austria, August 16-22, 1992.

SHANKER, S.G. 1988. (editor), *Gödel's theorem in focus*, Beckenham/New York, Croom Helm.

WANG, Hao 1974. *From mathematics to philosophy*, New York, Humanities Press.

—. 1987. *Reflections on Kurt Gödel*, Cambridge, Mass., MIT Press.