THE BULLETIN OF SYMBOLIC LOGIC
Volume 11, Number 2, June 2005

ON GÖDEL'S WAY IN: THE INFLUENCE OF RUDOLF CARNAP

WARREN GOLDFARB

The philosopher Rudolf Carnap (1891–1970), although not himself an originator of mathematical advances in logic, was much involved in the development of the subject. He was the most important and deepest philosopher of the Vienna Circle of logical positivists, or, to use the label Carnap later preferred, logical empiricists. It was Carnap who gave the most fully developed and sophisticated form to the linguistic doctrine of logical and mathematical truth: the view that the truths of mathematics and logic do not describe some Platonistic realm, but rather are artifacts of the way we establish a language in which to speak of the factual, empirical world, fall-outs of the representational capacity of language. (This view has its roots in Wittgenstein's *Tractatus*, but Wittgenstein's remarks on mathematics beyond first-order logic are notoriously sparse and cryptic.) Carnap was also the thinker who, after Russell, most emphasized the importance of modern logic, and the distinctive advances it enables in the foundations of mathematics, to contemporary philosophy. It was through Carnap's urgings, abetted by Hans Hahn, once Carnap arrived in Vienna as *Privatdozent* in philosophy in 1926, that the Vienna Circle began to take logic seriously and that positivist philosophy began to grapple with the question of how an account of mathematics compatible with empiricism can be given (see *Goldfarb 1996*).

A particular facet of Carnap's influence is not widely appreciated: it was Carnap who introduced Kurt Gödel to logic, in the serious sense. Although Gödel seems to have attended a course of Schlick's on philosophy of mathematics in 1925–26, his second year at the University, he did not at that time pursue logic further, nor did the seminar leave much of a trace on him. In the early summer of 1928, however, Carnap gave two lectures to the Circle which Gödel attended, or so I surmise. At these occasions, Carnap presented material from his manuscript treatise, *Untersuchungen zur allgemeinen Axiomatik*, that is, "Investigations into general axiomatics", which dealt with questions of consistency, completeness and categoricity. Carnap later circulated this material to various people including Gödel.

Received February 2, 2005; accepted February 9, 2005.
I am grateful to Steven Awodey, John Baldwin, Solomon Feferman, Akihiro Kanamori, and Stephen Simpson for helpful suggestions, and to Enzo De Pellegrin for research and editorial assistance.

1079-8986/05/1102-0006/$1.90

Gödel's serious interest in logic dates from that time. Subsequently he began the systematic reading in logic that brought him to the frontiers of what was then known, as his library request-slips show (*Dawson 1997*, p. 53). So, for example, in a letter to his fellow student Herbert Feigl, later a distinguished philosopher of science in America, he reports on what he did with his 1928 summer vacation:

> I myself was in Brünn the whole time and among other things read a part of *Principia mathematica*, about which, however, I was less enthusiastic [*begeistert*] than I had expected from its reputation. (letter to Feigl, 24 September 1928; *Gödel 2003*, p. 403)

That he was put onto *Principia* is no doubt also Carnap's influence; Carnap's concern with details of *Principia* far outstripped what would have been common among mathematicians by that time. It is, of course, a simplified version of the system of *Principia* that appears in Gödel's 1931 incompleteness paper.

Hilbert and Ackermann's *Grundzüge der theoretischen Logik*, published in 1928, was another of the first books Gödel looked at in logic. He found his dissertation problem in it: the question of the completeness of first-order logic is explicitly formulated for the first time in the book, on p. 68. In 1975, the sociologist Burke Grandjean sent a questionnaire to Gödel, in which he asked, "When did you become interested in the problem … of the completeness of logic and mathematics?" and "Are there any influences you would single out as especially important in this regard?" Gödel's answers were: "1928"; and "Hilbert Ackermann: Introduction to math Logic, Carnap: Lectures of math Logic" (*Gödel 2003*, p. 447).

Now, Carnap's logical work did not influence Gödel in any mathematical or technical way: there is in Carnap's material no mathematical idea that could be exploited for serious results. What Gödel learned from Carnap concerned the concepts that needed to be investigated. The peculiarities of the way Carnap framed those concepts motivated Gödel concerning both the problems he set himself to investigate and the correct formulation of the concepts.

Carnap's focus in the *Untersuchungen* was on properties of axiom systems such as consistency and completeness.[1] The subtitle of the first volume of the manuscript, and the title of Carnap's lectures, was *Metalogik*. But in fact Carnap's work was not at all metalogical, in our sense. For Carnap worked in a Russellian tradition, that is, with a conception of logic that can be called "universalist" (the term stems from *van Heijenoort 1967*). All notions were to be defined *within* the logical system. Consider the conjunction F of axioms of a system (Carnap's treatment is limited to finitely axiomatized theories). F contains various nonlogical vocabulary; let us call $F(R)$ what we obtain from F by replacing these with variables (of the appropriate types). Then

[1] Carnap never published this manuscript; it appeared only recently as *Carnap 2000*. In my account of this work in the next few paragraphs, I draw heavily on *Awodey and Carus 2001*.

the assertion that the axioms have a model (are satisfiable) is expressed in the system by

$$(\exists R)F(R).$$

The assertion that an assertion G follows from the axioms is expressed as

$$(\forall R)(F(R) \supset G(R)),$$

that is, in the terminology of *Russell 1903*, $F(R)$ formally implies $G(R)$. The assertion that the axioms are inconsistent is then the assertion that a contradiction follows from them:

$$(\exists G)(\forall R)(F(R) \supset G(R).{\sim}G(R)).$$

The quantification "$(\exists G)$" is to be construed here not as over formulas, but rather over propositional functions (i.e., properties) of a higher order.

Note how simple it is to show, once the concepts are defined in this way, that if the axioms are consistent then they have a model. For assume they do not have a model: that is, $(\forall R){\sim}F(R)$. Then $(\forall R)(F(R) \supset H)$ for any H, and in particular, for some H of the form $G(R).{\sim}G(R)$. Hence the axioms are inconsistent.

So here we have a trivial proof of a claim that verbally sounds exactly like a completeness theorem: an axiom system has a model if it is consistent. Of course, it is not really such a theorem; in particular the definition of consequence that is used to frame that of consistency does not accord with what is ordinarily meant by a deductive consequence of the axiom system.

The opening of Gödel's doctoral dissertation, completed in July 1929, reacts to this pseudo-completeness theorem. He starts by being insistently explicit on what notion of consequence is at issue in the dissertation. In the second sentence, he frames his theorem thus: every valid formula of first-order logic "can be derived from the axioms by means of a finite sequence of formal inferences." Two sentences later, after reformulating the theorem as about consistent axiom systems, he elaborates, "Here 'consistent' means that no contradiction can be derived by means of finitely many formal inferences" (*Gödel 1986*, p. 61).

On the next page, Gödel gives an explicit formulation of what Carnap's argument shows, in a way that makes apparent its difference from completeness.

> If we replace the notion of logical consequence (that is, of being formally provable in finitely many steps) by implication in Russell's sense, more precisely, by *formal* implication, where the variables are the primitive notions of the axiom system in question, then the existence of a model for a consistent axiom system (now taken to mean one that *implies* no contradiction) follows from the fact that a false proposition implies any other, hence also every contradiction (whence the assertion follows at once by indirect argument).

> [fn:] This seems to have been noted for the first time by R. Carnap in a hitherto unpublished work, which he was kind enough to put at my disposal in a manuscript form.

Of course, the subtext here is that in defining the notions as he did, Carnap trivialized the problem. But on the surface, Gödel is most polite in giving Carnap credit.

A more serious criticism is contained in the sentences just before this paragraph, although that it pertains to Carnap is disguised. Indeed, the point of those sentences has seemed to many readers to be somewhat obscure (in their Introductory Note to the dissertation, *Gödel 1986*, p. 49, Dreben and van Heijenoort call the remarks "somewhat misleading").

> But one might perhaps think that the existence of the notions introduced through an axiom system is to be defined outright by the consistency of the axioms and that, therefore, a proof [of completeness] is to be rejected out of hand. This definition ... however, manifestly presupposes the axiom that every mathematical problem is solvable. Or, more precisely, it presupposes that we cannot prove the unsolvability of any problem. For, if the unsolvability of some problem (in the domain of real numbers, say) were proved, then, from the definition above, there would follow the existence of two non-isomorphic realizations of the axiom system for the real numbers, while on the other hand we can prove the isomorphism of any two realizations. We cannot at all exclude out of hand, however, a proof of the unsolvability of a problem if we observe that what is at issue here is only unsolvability by certain precisely stated formal means of inference. For all the notions that are considered here (provable, consistent, and so on) have an exact meaning only when we have precisely delimited the means of inference that are admitted.

The point of these remarks becomes clear, I think, once we understand two further concepts that are central in Carnap's logical investigations, namely, *Entscheidungsdefinitheit* and monomorphicity. The former word is usually translated "syntactic completeness"; but that would be inappropriate here, since, in line with his universalist view, Carnap does not define the notion in a metamathematical way. Instead it is defined as

$$(\forall G)(G \text{ is a consequence of } F \text{ or } {\sim}G \text{ is a consequence of } F),$$

where consequence is taken in Carnap's sense, that is, as formal implication:

$$(\forall G)[(\forall R)(F(R) \supset G(R)) \vee (\forall R)(F(R) \supset {\sim}G(R))].$$

Monomorphicity corresponds to what we call categoricity, and is defined roughly thus:

$$(\forall R)(\forall R')(F(R).F(R') \supset R \cong R').$$

Carnap claims to prove in his manuscript that a theory is *entscheidungsdefinit* if and only if it is monomorphic. When Carnap's notions are replaced with their modern correlates, we obtain a biconditional one direction of which—from syntactic completeness to categoricity—is simply false, since syntactic completeness yields only the elementary equivalence of models and no more; and the other direction of which is nearly true, in the first-order case, needing only to be amplified by cardinality considerations. In contrast, the direction from *Entscheidungsdefinitheit* to monomorphicity is straightforward; this illustrates what it is to work inside a universalist conception. For if F is not monomorphic, then there are R and R' such that $F(R)$ and $F(R')$, but R and R' are not isomorphic. Consider the property of being isomorphic to R'; that is precisely a G for which the claim of *Entscheidungsdefinitheit* will fail, since it is neither the case that every R such that $F(R)$ is isomorphic to R' nor is it the case that every R such that $F(R)$ is not isromorphic to R'. The point is that the quantifier $(\forall G)$ in the definition of *entscheidungsdefinit* ranges over properties, with no limitation to those that are expressible in a fixed vocabulary of a formal theory. (In fact, Carnap takes the quantifier to range only over properties that respect isomorphism, that is, properties G such that $(\forall R)(\forall S)(G(R).R \cong S \supset G(S))$.)

However, the direction of Carnap's claim relevant to Gödel's remark is from monomorphicity to *Entscheidungsdefinitheit*. Again, the proof is nearly trivial. If the axiom system F is not *entscheidungsdefinit* then, for some G, both G and $\sim G$ fail to be consequences of F; by Carnap's definition, this means that both

$$(\exists R)(F(R).G(R))$$

and

$$(\exists R)(F(R).{\sim}G(R)).$$

Clearly these two realizations cannot be isomorphic; so the system fails to be monomorphic.

Now it can be shown in the logical theory within which Carnap works (the theory of types) that Peano arithmetic is monomorphic. Hence Carnap's claim would imply that Peano arithmetic is *entscheidungsdefinit*. A similar remark would hold for the axiom system for the real numbers to which Gödel alludes.

Gödel's criticism is directed on the surface against those who define mathematical existence in terms of consistency. Carnap does not do this; but, as we have seen, given his definitions existence follows from consistency in one trivial step. Moreover, it follows without any attention to the particularities of the logical system within which the notions are framed, e.g., whether they are first-order or higher-order. Hence, Carnap's procedure is little different from defining existence as consistency. And then Gödel's point is clear: there may be cases in which we can prove all realizations of an axiom system are isomorphic, yet this *does not* settle whether all questions are formally decidable from the axioms. Thus, Carnap's *Entscheidungsdefinitheit* does *not* capture the notion of formal decidability of a problem. That is what Gödel emphasizes in the last sentence cited above.

It should be emphasized that Gödel's argument here is meant to be a general one. He is arguing that it is a mistake to identify mathematical existence and consistency *in general*, that is, across a range of logical systems, including higher-order ones. For if this identification is made in general, it is made for logics in which the categoricity of various axiom systems, e.g., Peano arithmetic or the theory of the real numbers, can be proved. In those cases, the identification immediately yields the conclusion that one can not show that any question in those systems is formally undecidable, because to show formal undecidability of F is precisely to show that both the axioms conjoined with F and the axioms conjoined with $\sim F$ are consistent, which (by the assumed identification) would show the existence of non-isomorphic realizations of the axioms.[2]

Of course Gödel's remark is also a foreshadowing of the incompleteness result, which he obtained the following year. In fact, Gödel's first public announcement of incompleteness emphasizes precisely the same point. (Carnap was among the first persons to whom Gödel mentioned his result, in a private conversation on 26 August 1930. How he explained it on that occasion, or in a subsequent conversation with Carnap on 29 August, is not preserved, but he did talk about the method of arithmetization.) The setting for the public announcement was the September 1930 Epistemology of the Exact Sciences conference in Königsberg, a meeting that brought together positivistically inclined philosophers and logicians from Poland, Germany, and Austria. The main topic of Gödel's talk (*Gödel *1930c*) was his completeness result. After presenting it, he continued:

> I would furthermore like to call attention to an application that can be made of what has been proved to the general theory of axiom systems. It concerns the concepts "*entscheidungsdefinit*" and "monomorphic". . . . One would suspect that there is a close connection between these two concepts, yet up to now such a connection has eluded general formulation. . . . In view of the developments presented here it can now be shown that, for a special class of axiom systems, namely those whose axioms can be expressed in

[2]Thus it is a mistake to read the passage as being primarily concerned with first-order logic. Indeed, first-order logic cannot be the brunt of the argument, because (as Gödel realizes) his own proof of completeness shows that existence and consistency *can* be identified in this case. Hence there cannot be any counterexample to the identification, and *a fortiori*, no counterexample arising from his argument. His point for the first-order case is only to show that existence and consistency should not be identified *definitionally*: it requires real work to show their equivalence. And that point is argued by a consideration about logical systems generally, with the counterexample coming in the higher-order case.

> the restricted functional calculus, *Entscheidungsdefinitheit* always follows from monomorphicity ... If the completeness theorem could also be proved for the higher parts of logic (the extended functional calculus), then it would be shown in complete generality that *Entscheidungsdefinitheit* follows from monomorphicity; and since we know, for example, that the Peano axiom system is monomorphic, from that the solvability of every problem of arithmetic and analysis expressible in *Principia mathematica* would follow.
>
> Such an extension of the completeness theorem is, however, impossible, as I have recently proved; that is, there are mathematical problems which, though they can be expressed in *Principia mathematica*, cannot be solved by the logical devices of *Principia mathematica*. (*Gödel 1995*, pp. 26–29, translation emended)

The reference to Carnap could not be clearer, given Gödel's terminology and the expression "general theory of axiom systems." The point now is that Carnap's attempted theorem can be seen definitely to fail, once *entscheidungsdefinit* is defined in the appropriately metatheoretical way, as a direct consequence of incompleteness.[3]

In this material we can, I think, see a motivation for the incompleteness result that differs from characterizations Gödel gave later on. In the 1960s and early 1970s, Gödel often commented that what led him to his theorem was his recognition of a distinction between mathematical truth and provability, a distinction glossed over or even denied by both the Hilbert school and the positivists (or so Gödel alleged).[4] Nothing in what I have said goes against this altogether. But the remarks at the beginning of the dissertation and at the end of *Gödel *1930c* suggest more of a concern to underline the difference between consequence construed semantically (or universalistically) and syntactic consequence, that is, deducibility in a finite number of formal steps. Thus one need not impute back to the Gödel of 1930 the full-fledged Platonism of his later years as a motivator for the theorem.[5] Moreover, Gödel's

[3] In these remarks, Gödel seems unaware that an axiom system "expressed in the restricted functional calculus" (that is, expressed in first-order quantification theory) is categorical only in the relateively trivial case that all its models are of one finite cardinality (and are all isomorphic). This fact follows at once from the upward and downward Löwenheim-Skolem Theorems. Gödel did know of the latter: indeed, in his dissertation he cites *Skolem 1920* and talks of "the well-known theorem named for him and Löwenheim" (*Gödel 1986*, p. 77). Hence it appears likely that at this time he did not know of the former.

[4] Gödel also said that he was led to the theorem by attempting to give a relative consistency proof of analysis, vis-à-vis arithmetic (see *Gödel 2003*, p. 10). This concern is perhaps reflected in his use in the dissertation of the example of the real numbers as potentially yielding unsolvable problems.

[5] There are two other hints in his early writing that Gödel was not as full-fledgedly a Platonist in the early 1930s as he later became. In *Gödel *1933o*, a lecture given in Cambridge, Massachusetts, Gödel finds problematic the use of impredicative definitions for specifying

concern to show the incompleteness of higher-order logic provides a reason for his presenting the result in the 1931 paper as one about the theory of types, rather than as about axiomatic theories of number theory, formulated within first-order logic, as became the common expository practice, starting with Gödel's own Princeton lectures of 1934.

That, then, is Carnap's influence on Gödel. After incompleteness, the influence goes in the other direction. Carnap had started to recognize the flaws in his conceptualization in early 1930, under the urging of Tarski. His *Untersuchungen* was abandoned by April of that year, before Gödel told him of incompleteness. But incompleteness, and the technique of arithmetization, gave him the great spur for his principal project, *The Logical Syntax of Language*, published in 1934.

> My way of thinking was influenced chiefly by the investigations of Hilbert and Tarski in metamathematics ... I often talked with Gödel about these problems. In August 1930 he explained to me his new method of correlating numbers with signs and expressions. Thus a theory of the forms of expressions could be formulated with the help of concepts of arithmetic. (*Carnap 1963*, p. 53)

Arithmetization was important to Carnap, because the ability to frame syntax within a clearly unobjectionable arithmetical language answered the doubts, stemming from Wittgenstein and shared by some members of the Vienna Circle, that the logical structure of language could not properly be described at all.

Much of *Logical Syntax of Language* is a response to the challenge Carnap took to be posed by Gödelian incompleteness.[6] Incompleteness shows that the notion of mathematical truth cannot be captured by notions based on formal derivability. Hence if mathematical truth is to be an artifact of language, languages could not be determined solely by the deductive links they contain. Carnap sought definitions that go beyond deductive ones, could be thought of as consitutive of languages, and would enable him to capture mathematical truth, while using only the resources he considered "syntactic" (a word he used in a wider sense than it currently has). In this project Gödel was of technical assistance as well: he showed Carnap his

classes, and concludes his discussion of foundational issues by saying, "The result of the preceding discussion is that our axioms, if interpreted as meaningful statements, necessarily presuppose a kind of Platonism, which cannot satisfy any critical mind and which does not even produce the conviction that they are consistent" (*Gödel 1995*, p. 50; cf. Feferman's Introductory Note, *ibid.*, pp. 39–41). And in a letter to Carnap of 28 November, 1932, about the definition of mathematical truth (in Carnap's terminology, "analyticity") Gödel says, "I believe that the interest of this definition does not lie in a clarification of the concept 'analytic', since one employs in it the concepts 'arbitrary sets', etc., which *are just as problematic*" (*Gödel 2003*, p. 357; my emphasis).

[6]In the book, Carnap presents a detailed proof of the First Incompleteness Theorem; he was the first to formulate this by isolating a general Fixed Point (Diagonal) Lemma.

attempt to obtain mathematical truth by induction on syntactic form would fail for higher order quantifiers, due to impredicativity; and he pointed him in the direction of the appropriate definition. (See *Gödel 2003*, pp. 347–357.) As a result, Carnap wound up with a specification of mathematical truth essentially equivalent to Tarski's, at about the time Tarski was first publishing his own treatment. In any case, Carnap's *Logical Syntax* is in both motivation and technique inconceivable without Gödel.

REFERENCES

[2001] Steven Awodey and André W. Carus, *Carnap, completeness and categoricity: the* Gabelbarkeitssatz *of 1928*, ***Erkenntnis***, vol. 54 (2001), pp. 145–171.

[1934] Rudolf Carnap, ***Logische Syntax der Sprache***, Springer, Vienna, 1934.

[1963] ———, *Intellectual autobiography*, ***The philosophy of Rudolf Carnap*** (Paul A. Schlipp, editor), Library of Living Philosophers, vol. 11, Open Court, La Salle; Cambridge University Press, London, 1963, pp. 3–84.

[2000] ———, ***Untersuchungen zur allgemeinen Axiomatik***, Wissenschaftliche Buchgesellschaft, Darmstadt, 2000, (Thomas Bonk and Jesus Mosterin, editors).

[1997] John W. Dawson, Jr., ***Logical dilemmas: The life and work of Kurt Gödel***, A K Peters, Ltd., Wellesley, Mass, 1997.

[1929] Kurt Gödel, ***Über die Vollständigkeit des Logikkalküls***, doctoral dissertation, 1929, in *Gödel 1986*, pp. 60–101.

[*1930c] ———, *Vortrag über Vollständigkeit des Funktionenkalküls*, *1930c, in *Gödel 1995*, pp. 16–29.

[1931] ———, *Über formal unentscheidbare Sätze der* Principia mathematica *und verwandter Systeme I*, ***Monatshefte für Mathematik und Physik***, vol. 38 (1931), pp. 173–198, in *Gödel 1986*, pp. 144–195.

[*1933o] ———, *The present situation in the foundations of mathematics*, 1933, in *Gödel 1995*, pp. 45–53.

[1986] ———, ***Collected works. Vol. I: Publications, 1929–1936***, Oxford University Press, New York and Oxford, 1986, (Solomon Feferman et al., editors).

[1995] ———, ***Collected works. Vol. III: Unpublished essays and lectures***, Oxford University Press, New York and Oxford, 1995, (Solomon Feferman et al., editors).

[2003] ———, ***Collected works. Vol. IV: Correspondence A–G***, Oxford University Press, New York and Oxford, 2003, (Solomon Feferman et al., editors).

[1996] Warren Goldfarb, *The philosophy of mathematics in early positivism*, ***Origins of logical empiricism*** (R. N. Giere and A. W. Richardson, editors), University of Minnesota Press, Minneapolis, 1996, pp. 213–230.

[1928] David Hilbert and Wilhelm Ackermann, ***Grundzüge der theoretischen Logik***, Springer, Berlin, 1928.

[1903] Bertrand Russell, ***The principles of mathematics***, Allen and Unwin, London, 1903.

[1920] Thoralf Skolem, *Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theoreme über dichte Mengen*, ***Skrifter utgit av Videnskapsselskapet i Kristiania, I, Matematisknaturvidenskapelig klasse***, no. 4, 1920, pp. 1–36.

[1967] Jean van Heijenoort, *Logic as calculus and logic as language*, ***Boston studies in the philosophy of science***, vol. 3 (1967), pp. 440–446.

DEPARTMENT OF PHILOSOPHY, HARVARD UNIVERSITY, CAMBRIDGE, MA 02138, USA
E-mail: goldfarb@fas.harvard.edu