8

FORMAL-SYNTACTICAL THINKING AND THE STRUCTURE OF THE WORLD

Paul M. Livingston

It is distinctive of 'analytic' philosophy in the twentieth century to see traditional problems of the thinking-being relationship as, essentially, problems of the bearing of *language* on reality: how do the signs and structures of our finitely formulated and rationally intelligible language come to describe and characterize the facts and objects of the world, as such and as a whole? As I shall argue here, the internal development of this problematic over the course of the tradition establishes not only the specific forms in which this bearing can be rationally comprehended, but also the essential formal limits to which this comprehension is subject. The demonstration of these formally necessary limits then essentially henceforth marks the broader problematic of the thinking-being relationship, most decisively where linguistic reflection considers the implications of its own *position* within the totality of the world that it thinks.

One of the most decisive innovations of analytic philosophy in its initial stages was the prospect, suggested by Wittgenstein but developed most prominently by Carnap, of understanding rational thinking about the world as having the rule-governed formal structure of a language: the finitely tractable 'logical syntax' of its well-specified rules of symbolic combination and transformation. A language is, on this conception, essentially a calculus: that is, a structure of signs governed by logical or syntactical rules for formation and transformation that are *formal* in the sense that they are empty of material content and independent of contingent facts. It is only as such that the signs of a language are, on this conception, capable at all of determinate and non-contradictory application to the description of the world, as *all that is the case*. But the comprehensive perspective of formal-syntactical thought also necessarily raises—as I shall argue here—the essential problem of the position from which this thought *itself* takes place. This is the reflexive problem of the position from which it is possible for formal thinking to (1) propound or discover the formal structure of a language and, at the same time (2) ensure its application to the facts and truths of the world as a whole.

As the formal results of Gödel and Tarski would soon bear out, this reflexive problem has deep implications for the possible coherence of the project of a comprehensive formal-logical thinking of the structure of the world. In particular, the formalization of the reflexive problem itself, as the problem of the possibility of capturing in a finite language the totality of its own deductive procedures, demonstrates an essential formal *undecidability* inherent in any envisioned application of such a finitely comprehensible calculus to the unlimited totality of the world. This situation is, most directly, a consequence of the dilemma propounded by Gödel's two 'incompleteness' theorems: namely, that a formally structured language for mathematics cannot without contradiction be seen as, simultaneously, both *consistent* and *complete* in its capacity to capture mathematical truths. As Gödel himself argues in a 1951 lecture and several drafts (written between 1953 and 1959) of an unpublished paper critical of Carnap's formal-syntactical project, the consequences of this situation ultimately demand the rejection of that project, along with the joint assumption of the possible consistency and completeness of a formally specified language that it relies upon. As a result, formally and finitely comprehensible thought is brought, as I argue here, to witness the further dilemma of its own incapacity to grasp the total structure of the world by rational means, or the actual undecidability of that structure itself.

*

The idea of 'logical syntax' or 'grammar' as a system of rules governing the correct use of signs in a symbolic language is introduced by Wittgenstein at *Tractatus Logico-Philosophicus* 3.325. In that context, the suggestion responds to a characteristic danger to which we are, according to Wittgenstein, regularly prone in the unreflective use of our everyday language: since superficially similar signs are often used to signify in what are in fact importantly different ways, we are recurrently led to confusions of a fundamental kind, of which he says (in the preceding remark) philosophy is full. Given this, a clarifying logic of signs, which would aim to coordinate each sign to exactly one use, will be necessary to eliminate this kind of error and evince the real underlying structure of our language in its referential or descriptive relation to the world overall. Along with other main theses of the *Tractatus*—for example, that of propositions as the sensible expressions of thoughts (1921: 3.1), thoughts as logical pictures of facts (3), and pictures as picturing by virtue of their form, and thus *independently* of what is true or false (1921: 2.2, 2.22)—this suggests the decisive requirement that Wittgenstein now places on the establishment of logical syntax. This is the requirement that the rules for the use of signs overall be formally *empty*: only the syntactical description of the possible expressions themselves must be mentioned in the establishment of logical-syntactical rules, and it must never be necessary to refer to their referential or descriptive meanings (1921: 3.33). This emptiness of logical syntax then further implies the capacity of a logically syntactically structured language to capture the *global* structure of the facts of the world, or of all that is the case: to present [*stellen … dar*], as Wittgenstein says in 6.124, the world's 'scaffolding' [*Gerüst*] through the identification of those combinations of signs which, as tautologies, say nothing (1921: 6.11) and thereby, through the fact of their being tautologies, show [*zeigt*] the formal or logical properties of language and the world overall. (1921: 6.12) The formal study of the rules governing possible signs in their empty and arbitrary possibilities of formation and transformation thereby becomes, in a methodologically important sense, the study of the overall logical structure of the facts and phenomena of the world. And the definition and description of these possible rules is the formally empty display, in tautologies, of the logic of the facts of the world as such, and as a whole.

A further consequence of this conception which Wittgenstein draws is that the logic of the world cannot be presented *within* the world. Descriptive propositions, although they can represent all of reality, cannot represent the logical form that they must have in common with it as a whole (1921: 4.12). It is thus impossible for us to express the logical form of reality in assertible propositions, and this structure is rather to be *shown* by the discernment of logical relations among propositions and the recognition of logical principles as tautologies empty of content (1921: 4.121–22, 6.112–12). Similarly, the equations of mathematics are not propositions with content, and they express no thoughts; rather, they *show* the logic of the world by showing, without reference to facts, what expressions can be substituted for one another (1921: 6.22–24). In this way, the clarification of the tautologies of logic and the equations of mathematics, by clarifying the underlying formal rules of use, demonstrate the adequacy of signs to the world as a whole.

From the beginning of the Vienna Circle's discussions, the methodology of the clarification of the logic of language was a key component in the overall project of a 'scientific' philosophy dedicated to the logical and epistemological structure of scientific knowledge overall. The conception that the overall logic of science is to be understood, specifically, as 'nothing other than the logical syntax of the language of science' became the guiding methodological idea of Carnap's project in his 1934 *The Logical Syntax of Language*. Earlier, in his 1928 *The Logical Structure of the World*, Carnap had sought to describe the unitary totality of scientific facts by means of an overall logical analysis of the underlying structure of the concepts of science and their relations, a so-called 'constructional system' for the world as a whole. By 1931, Carnap had been convinced by Otto Neurath of the thesis and project of *physicalism*: that (as the two philosophers understood the claim) the unity of science, and thus of the world, can be understood to correspond to the universality of a single logically structured language, the so-called 'physical' one. In *The Logical Syntax*, these methodological conceptions combined with Wittgenstein's idea to produce the project of the analysis of the logical syntax of a language. Such an analysis, Carnap says in introducing the project, will take the form of a 'systematic statement of the formal rules' governing the formation and logical transformation of its expressions, together with a statement of the consequences of those rules (Carnap 1934: 1). As in the *Tractatus*, the rules are to be understood as purely 'formal' in that no reference is to be made in stating them either to the referential meaning of any individual signs or to the senses of the expressions as wholes. Rather, the rules are simply to be rules for the combination of signs into certain initial combinations (the so-called 'formation rules') and for the transformation of sequences of signs into other sequences (the so-called 'transformation rules'). Languages themselves are treated as calculi, in the sense that they are systems of 'conventions or rules' of this kind, and the practice of logical syntax itself is then nothing other than the 'construction and manipulation' of such calculi (Carnap 1934: 4).

As Carnap notes, the formal method here invoked bears close parallels to Hilbert's *formalist* program in the philosophy of mathematics (Carnap 1934: 9). On this program, claims about the infinite are to be replaced by formalization in a finite syntax, thereby gaining application to the statement of general mathematical truths and truths ranging over infinite domains; all that is required for the coherence of this statement is that the relevant systems can be shown to maintain consistency. For Carnap, mathematics is itself logical syntax, in the sense that the carefully constructed languages he envisions contain expressions (at least) for numbers and their relations, along with well-defined rules for the derivation or proof of arithmetical statements in general from formal-syntactic definitions (Carnap 1934: secs. 21, 22, 28, and 32).

Although he agrees with Wittgenstein in holding that syntactic rules are empty of content, Carnap differs sharply from Wittgenstein in adding to this the thesis of our *complete freedom* in propounding them. For Carnap in *Syntax*, rather than being necessary structures of a unitary language, formal-logical systems are systems of arbitrarily adoptable convention. Each such system defines a language, but there is nothing to constrain the freedom of the logician or philosopher in propounding and exploring the consequences of such systems. In *Syntax*, Carnap formulates this conventionalist position by means of the overall adoption of what he terms a 'Principle of Tolerance'. According to this principle, the constructional forms and the rules of transformation can be chosen completely arbitrarily. These arbitrarily chosen rules will then determine what are subsequently to be understood as the 'meanings' of the basic logical signs, rather than conversely, and what had seemed to be the study of logical necessities or formally binding truths can cede to the free and self-conscious exploration of the variety of syntactical systems (Carnap 1934: xv, 52).

If it is thus said to be possible for the philosopher or logician to describe or stipulate arbitrary sets of rules of syntax in this way, the question arises of how these rules can themselves be expressed. Is it necessary that the rules governing a specific language be presented in another, different (and perhaps 'stronger') language, or is it possible that the rules for a language be represented wholly within it, itself? The first alternative would seem to require, for the complete explication of all the relevant rules, an infinite hierarchy of languages; whereas the second raises the spectre, as Carnap notes, of the paradoxes of self-reference and self-inclusion. Nevertheless, Carnap opts for this latter alternative as a general matter, citing specifically the technique for representing formulas and sentences of a language within that language itself that had recently been developed by Gödel in arriving at his two 'incompleteness' theorems. By means of this technique, the so-called 'arithmetization of syntax', the expressions can be represented by numbers and the syntax by arithmetical relations of these numbers, so that, provided the system has the expressive power to represent arithmetic, it can also represent its own essential syntactical rules (including the rules of formation and transformation) as a whole. This technique and possibility, as Carnap emphasizes at several points in *Syntax*, appears to provide a dramatic alternative to Wittgenstein's view, according to which logical syntax cannot be represented but only shown or indicated at the limits of the world of facts. For it provides that, both with respect to artificially constructed languages and the natural ones whose structure the clarification of logical syntax may help to illuminate, the structure of conventional rules constitutive of a language *can*, in general, be represented in that language itself. At the same time, however, this opens the significant question of the formal consequences of this kind of linguistic reflexivity, or of the implications of the position of formal-syntactic thought *within* the scope of the language or languages it, itself, describes.

This problem about the extent and limits of a language's capacity to represent itself within itself would, at any rate, soon have decisive implications for analytic philosophers' thinking about syntax, semantics, and truth. In 1936, two years after the publication of *Syntax*, Alfred Tarski considered in his 'The Concept of Truth in Formalized Languages' the possibility of a definition of *truth* for syntactically well-specified languages in general. Such a definition, for Tarski, would take the form of a systematic definition of a predicate holding of just those sentences that are (intuitively) true; that is, just those sentences which say what is the case. In other words, it would have as consequences each of the sentences of the convention or schema T:

T: 'S' is true *iff* S

A definition of truth that is adequate in this sense will be possible, wherever a language is capable of making reference, in general, to its own sentences, for example by means of quotation marks (as earlier) or by means of a suitable arithmetization device. However, as Tarski argues in 'The Concept of Truth', if the language is both 'universal' in its expressive resources—able, that is, to make such general reference to *any* of its own sentences—and understood as capable of formulating its *own* truth predicate, the formal schema will then necessarily demonstrate the *inconsistency* of the language in question (Tarski 1936, 164–65). For it will be possible to produce a sentence asserting its own falsehood (this is a version of the classic paradox of the Liar) and then the consequence of the T-schema will be that that sentence is true if false, and vice versa. In 'The Concept of Truth', Tarski draws the conclusion that, whereas natural languages such as English which evidently include their own truth-predicates are very likely to be inconsistent, and therefore incapable of consistent overall syntactic analysis, the truth-definition for a (presumably consistent) formal language can nevertheless be carried out, provided only that the language in question is *not* expressively universal in the above sense, and that the truth-definition is, accordingly, not fully definable within it, but only within *another* language of greater expressive power. The full definition of truth for *this* language would then, however, evidently require to be carried out in another, stronger one; and so forth indefinitely.

This result would play an essential role in convincing philosophers of the ultimate untenability of any *purely syntactical* conception of truth-conditional meaning in a language overall. Instead, as Tarski already suggests in 1936 and clarifies in the 1944 article 'The Semantic Conception of Truth', truth is to be understood not as a purely syntactic but rather a partially *semantic* notion, in the sense that its formalization depends essentially on claims about the reference or denotation of various terms, and more broadly on the relationship of the language in question to some specifiable (and thus delimited) set of objects external to itself. Familiarly, this 'semantic' conception is one of the central roots of the pursuit of formal semantics as (what is later called) 'model theory', whereby in addition to the syntactic rules for a language, properties of the language are considered in their representative or referential relationship to objects and relations in such domains.

In this way, the problem that is effectively posed in Tarski's analysis by the inability of the formal syntax of a 'universal' language to represent itself led in part to the widespread abandonment by analytic philosophers of the straightforward project of a logical-syntactical analysis, at least with respect to the structure of natural languages of 'universal' expressive scope. Nevertheless, as I shall argue in the final section of this paper, equally significant aspects of the formal-syntactical conception nevertheless remain, and continue to characterize central orientational and methodological commitments of central analytic projects up to the present. In particular, a wide variety of such projects they still require that the illumination of underlying and finitely specifiable rules be able to clarify the structural basis of the referential capacities and truth-evaluability of sentences in the language as a whole.[1] As such, they require that the consequences of the specifiable rules follow from these *logically* and with respect to the world as a whole. In this respect, as I shall argue, they also pose the *positional* question of the place from which these consequences are themselves drawn; and it is the unavoidability of this question which, most directly and comprehensively, exposes constructional thinking to the consequence of the ultimate *undecidability* of the thinking-being relation it aims to formulate.

*

The two 'incompleteness' theorems that Gödel demonstrated in 1931 bear most directly on finitary formal systems for mathematics, of the type pursued by Hilbert in his formalist program.[2] They show (on a relatively uncontroversial formulation) that for any such system capable of representing some portion of arithmetic, *if* the system is consistent: (1) there will be statements P in the language of the system such that the system does not suffice to prove either P or ~P, and which are thus said to be 'undecidable'[3]; and (2) the system cannot prove a statement of its own consistency. If supplemented with the claim, with respect a given 'undecidable' statement, that this statement has a determinate truth value (either true or false), then the first result bears out the claim that the system fails to prove at least one truth, and is therefore (in this sense) 'incomplete', again on the assumption that it is consistent at all. Together, the two theorems were widely taken to defeat Hilbert's formalist project of establishing finitary formal systems for mathematics: for they were taken to show that any such system will be unable, if sound (i.e., if it proves no falsehoods) to capture *all* of the statements that *we* can nevertheless see to be mathematically true, and further that no such finitary system is able to verify by its own means its own consistency (and hence soundness) at all.

On the basis of the two theorems, Gödel challenged Carnap's conception of logical syntax in the 1951 Gibbs lecture 'Some Basic Theorems on the Foundations of Mathematics and Their Implications', and six drafts, written between 1953 and 1959, of an article entitled 'Is Mathematics Syntax of Language?' In the 1951 lecture, Gödel first describes an overall dilemma that evidently characterizes the situation of human mathematical thinking, in view of the two theorems and the equivalence later established by Turing between the structure of formal systems and finite computational machines (i.e., 'Turing machines'). As a consequence of the second theorem in particular, which shows that no well-defined system of axioms and rules can prove a statement of its own consistency (and hence its own correctness or soundness, since a system is [presumably] sound only if consistent), it is contradictory for anyone to set up such a finite system and claim to know simultaneously *both* that the system is complete—sufficient to establish *all* mathematical truths—*and* that it is correct (Gödel 1951: 309). For the claimed insight into the correctness of the system, and hence its consistency, is, if real (due to second theorem) a mathematical insight that *cannot* be derived in the system itself; and so by that very insight, there is a truth that goes beyond what the system can prove, and thus the system is then not complete.

This leads Gödel to state the more general dilemma which then plausibly characterizes the situation of our own mathematical cognition relative to that of finite formal systems. Given that it has been proved that, for any such system, there will be arithmetical statements whose truth-value the specified system cannot decide, either our own powers of mathematical cognition to establish such truths (including the statements of the consistency of the relevant systems themselves) exceeds those of *any* such system, or they do not. In the first case, the human mind is capable of infinitely exceeding the capacities of any possible Turing machine, in that it can again and again see as evident ever-new mathematical truths that cannot be, as a whole, comprised by any finite rule. On the second alternative, however (on which the mind is after all equivalent to some particular formal system), there will then always be certain straightforward types of arithmetic statements that are *absolutely* undecidable: undecidable, that is, by any proof that the mind can conceive (Gödel 1951: 310). However, on either alternative, Gödel argues, some form of realism about mathematical objects and truths appears to follow, in that it is untenable to suppose (on either alternative) that mathematics is our 'free creation'. On the first alternative, whereby the mind's power to grasp certain mathematical truths exceeds that of any finite system, it is evidently the case that this power must then be conceived as essentially going beyond anything that we (as finite beings) can create. However, on the second alternative as well (that of absolutely unsolvable problems) it appears implausible that the truths are created by us, since if they were, in creating them we would necessarily know all of their properties; and in view of the absolute insolubility of the relevant problems (on this alternative), we do not.

Gödel now addresses specifically Carnap's conception of logical syntax as 'the most precise, and at the same time most radical, formulation' that has been given so far of the view that mathematical propositions are expressions solely of our own syntactic/linguistic conventions. On the view as Gödel sketches it, mathematical propositions are, as a whole, consequences of syntactic conventions that do not refer to any extralinguistic objects, and are thus analytic in the sense that they are empty of content and do not imply the truth or falsity of any factual proposition (Gödel 1951: 315–16).[4] As Gödel recognizes, and as Carnap has shown in *Logical Syntax* (following an earlier demonstration by Ramsey), the truth of standard sets of axioms for mathematical inference (for example the axioms of Peano Arithmetic) can indeed be shown to be derivable from suitably chosen purely syntactic rules in this sense. *However*—and here is the decisive objection—the derivation even of the basic axioms *from* the semantic rules *must itself make extensive use* of mathematical concepts and principles, in application to the syntax, that cannot themselves be known to be true *unless* the axioms and their consequences are *already* so known. Further, and even more decisively, a proof of the possibility of deriving the axioms from syntactical conventions (and hence of their purely tautological character) is at the same time a proof of the *consistency* of those axioms themselves; and, as a result of the second 'incompleteness' theorem, this proof is impossible using *only* those axioms themselves. The verification of the following of the axioms from the syntactic rules—and thus their emptiness and purely tautological status—must then require knowledge of mathematical truths that essentially goes beyond what can be derived from these axioms themselves. It follows, for any finitely specifiable set of syntactic rules, even the possibility of portraying those rules as the syntactic basis for a system of axioms requires knowledge of the truth of propositions that can be shown to go beyond any that can be derived from those rules. For this reason, Gödel concludes, 'there exists no rational justification of our precritical beliefs concerning the applicability and consistency of classical mathematics (nor even its undermost level, number theory) on the basis of a syntactical interpretation', (Gödel 1951: 318) and, at least with respect to mathematical propositions, this conception is decisively refuted.

In version III (the longest draft version) of 'Is Mathematics Syntax of Language?' Gödel repeats and extends the criticism, citing not only Carnap but also Hahn and Schlick as adherents to the conception according to which mathematics can be interpreted as (finitary) syntax of language. On any such conception, it will be the case, first, that the formal axioms and procedures of mathematics can be derived from purely syntactical rules; and secondly, that in the case of any conclusion about ascertainable empirical facts that were formerly justified in part by means of intuitive mathematical considerations, those considerations can be replaced by the consequences of purely syntactical rules. However, as Gödel argues, this will be possible *only* in the presence of a consistency proof for those rules. For example, from a formal-syntactic proof of the truth of Goldbach's conjecture (every even number > 2 is the sum of two primes) it would be possible to predict the empirical behavior of a certain calculating machine from the syntactic conventions, but *only* if they are known to be consistent; otherwise, as Gödel points out, they make *no* determinate prediction about the possible behavior of the machine at all (Gödel 1953: 339–40). In light of the second theorem, however, any possible proof of such consistency necessarily relies on mathematical principles that go beyond any that can be derived from those rules themselves. Thus, in order to justify the syntactical program, the whole program must be 'turned into its downright opposite' (Gödel 1953: 341–42) in the sense that, instead of specifying the conceptual meanings of mathematical terms by pointing to syntactical rules, it must instead make essential and extended use of these meanings to establish these very rules as consistent, and hence applicable, at all.

Finally, both here and in the 1951 lecture, Gödel points out another consequence of the issue of consistency which appears devastating, not only to Carnap's specific formulation of the logical-syntactical project but for the broader logical positivist strategy of separating propositions into the categories of the empirical-factual (or synthetic) and formal-logical (or analytic) at all. Since, under standard logical assumptions, a contradiction implies *all* sentences, in the absence of knowledge of the consistency of the formal-syntactic rules it will not be known that these rules do not have as consequences all sentences of the language, whether they be of 'formal' or 'empirical' character. It will thus be impossible to consider the sentences of the language to be separable, as a whole, into the two categories of those bearing empirical content and those, expressive only of the syntactical conventions, which lack it. Indeed, since the statements intended to capture syntactic rules cannot then be shown to be empty of content, it appears impossible to consider any set of them as capturing wholly and exclusively the meanings of terms in the language overall, and the underlying conception of a language as a structure or system of such purely syntactical rules is thereby defeated.

All of these considerations directed against the coherence of the logical-syntactical conception of mathematics turn on the proven inability of finitary systems of rules (in the sense of a finitary formal system or a Turing machine) to establish, by finite means, the truth or falsity of certain further mathematical statements (including the statement of the system's own consistency) whose truth would have to be known in order to establish any meaningful application of these rules themselves. We can put this consideration, in general terms, as that of the *undecidability* of the application of formal-syntactical rules beyond themselves: their inability, that is, to establish their *own* total application to the range of formal truths they are supposed to underlie, and to the overall distinction between formal and empirical truths itself.[5] If, as Gödel argues, this undecidability of application is a necessary feature of any finitely formulable system of rules, then it will follow that the logical syntax of any finitely specifiable language is undecidable in general. Every finitely specifiable language will contain infinitely many sentences that are undecidable by its own syntactical rules, and it will moreover be impossible to restrict the range of undecidability to those sentences that are meant to be empirical as opposed to those that are supposed to express purely formal truths. As a consequence, the formal structure purportedly underlying truth-conditional meaning in the language as a whole will itself be undecidable in this sense: any finitary specification of the syntactic rules will then leave this meaning radically undetermined, and the logical syntax project will fail in general.

As commentators have emphasized and as Gödel himself notes, to the extent that Carnap at the time of *Logical Syntax* has a response to these objections, it turns on the latter's invocation of *infinitary* means of proof and demonstration. In particular, in *Syntax,* Carnap (1934: 39) distinguishes between the kind of finitary method of deduction suggested by Hilbert, wherein a derivation is a finite series of sentences, and (what Carnap understands as) a broader method, which he calls the 'method of the consequence-series'. By contrast with finitary derivations, a deduction by the latter method may be a deduction from (simultaneously) a set of infinitely many sentences as premises, or from any finite number of such sets: a deduction carried out in this way will then not be 'definite' in each of its steps in the way that a finite derivation is, but will nevertheless establish (Carnap says) that its result follows from its (possibly infinitely many) premises as a 'consequence' (Carnap 1934: 39). Analytic sentences, or tautologies, are then defined as those that are 'consequences', in this sense, of the empty set of sentences of the language, and thus (also) of every sentence, and it is possible in this way, according to Carnap, to verify the tautological character of mathematical sentences in general, as well as confirm various broader properties of the specific formal languages investigated.

It is this appeal to the infinitary method of deduction that allows Carnap, in *Logical Syntax*, to avoid or re-situate what are normally seen as the problematic consequences of Gödel's theorems, of which Carnap was well aware at the time. For example, by means of the application of such an infinitary deduction rule, the consistency of the specific Language II constructed by Carnap can, he holds, be demonstrated, albeit by means that, as he says, go essentially 'beyond the resources at the disposal of Language II' itself (Carnap 1934: 129). Indeed, more generally, the demonstration of the non-contradictoriness of any language will, quite plausibly in the light of Gödel's results, require such means that go beyond its own (Carnap 1934: 134, 219). Thus, it is necessary, according to Carnap, to formulate any such demonstration for a particular language, S, in another language S_1, and so on; indeed, anticipating Tarski's results of two years later, Carnap here indicates that such a methodology will be necessary for any coherent treatment of a language's truth-predicate (Carnap 1934: 216). Similarly, since (by Gödel's first theorem), any language capable of expressing arithmetic, if consistent at all, will include sentences that are undemonstrable by its means, each such system is 'defective' in the sense that it cannot by its own means determine the truth-value of all arithmetical sentences formulable within it. But this does not imply, Carnap suggests, any actual indeterminacy, since each such sentence may indeed be decided in another system. In this sense, and for this reason, although 'there exists neither a language in which all arithmetical terms can be defined nor one in which all arithmetical sentences are resoluble [decidable]', it nevertheless remains the case that 'everything mathematical can be formalized' by means of the continual progress of the methods of demonstration through an infinite series of languages, each stronger than the last (Carnap 1934: 222).

Whereas Carnap thus seeks to immunize the logical-syntactic conception of the consequences of essential undecidability and the indemonstrability of consistency by appeal to such a hierarchy of languages and the non-finitary consequence rule, Gödel by contrast takes the appeal to such non-finitary means to be a *reduction* of the whole logical-syntactical program. For, as he points out, any such appeal that is sufficient to establish (so much as) the consistency of our logical-syntactical conventions, and hence their applicability to any sentences beyond themselves, will involve mathematical knowledge beyond that which can be established by their finitary application alone. It is not that Gödel himself disbelieves, in general, in the possibility of such non-finitary knowledge, including knowledge of the consistency of any set of conventions which we do in fact formulate, at a particular time, as capturing our practices of mathematical reasoning. This knowledge might be gained, for instance, by our being able to have, in each case, a non-syntactic (and non-methodical) *intuition* of the relevant truth; or it might be seen as grounded in our ability to perceive certain substantive facts, going beyond linguistic conventions, about the relationships of their constitutive concepts.[6] We might, indeed, be capable of, again and again, formulating the newly gained knowledge in finitary terms, for example, as new axioms continually introduced (albeit not into a *single* language, since the new formulation would produce a new language in each case). But what Gödel points out is that, if we appeal to this possibility of iterated reflection in general, we are no longer appealing to *syntactic* considerations in any real sense. Rather, we are here appealing to capacities for knowledge that must essentially exceed the operation of any finite structure or mechanism, and thus must (infinitely) exceed anything we can capture *completely* by means of any linguistic formulations that we can understand at all.[7] It follows that, if we can indeed represent the envisioned capacities to ourselves, we will not represent them as rational ones; and, if we do in fact possess them, we will not be able to make this possession rationally intelligible to ourselves.

We can put this consideration in vivid form by considering what is required, as a matter of epistemic position, of any knower who would in fact be capable of applying one of Carnap's purported infinitary rules of deduction. A simple form of such a rule is the so-called 'omega-rule', which 'instructs' its user to conclude $\forall x\phi(x)$ from the infinitely many distinct premises $\phi(1)$, $\phi(2)$, $\phi(3)$, and so on, for every natural number. The inference appears at first relatively unproblematic, but it is important to note that in order to form the basis for the application of the rule, the infinitely many premises must first be *independently* known: each one, that is, established 'on its own', without any evident rule uniting them all. The knowledge of the truth of these infinitely many premises cannot be, for example, the result of an application of the usual (finitely formulable) rule of mathematical induction: for if that rule were sufficient, the general conclusion could already be established on its basis, without any need to appeal to infinitary procedures. Rather, the knower who is able to apply the infinitary 'rule' must already be in a position to know the properties of each number without inferring these directly from the properties of any other: 'one by one', so to speak.

But what kind of knower could be in such a position? We can imagine such a knower, apparently, only as one who is capable of a kind of infinite survey of *all* natural numbers, or at any rate capable of performing infinitely many finite procedures at once. To credit a knower with *these* capacities is, however, evidently to credit her with the ability to know the properties of each natural number even *before* carrying out the (purported) infinite deduction; and to do this would be, presumably, to beg the question of the extent of her possible knowledge from the position she is envisioned as occupying. If, however, we cannot characterize a putative knower as applying an infinitary rule to gain mathematical knowledge unless we portray her as already possessing the relevant knowledge, then the appeal to the infinitary rule is idle, and cannot support the claim that there are rational procedures capable in general of yielding the relevant knowledge. But then, the only evident alternative is to portray the means we have of attaining this knowledge as consisting in something other than rational procedures, if we are capable of attaining it at all.[8]

*

As we have seen, if the considerations that Gödel introduces against the coherence of the formal-syntactical project are understood as posing the broader problem of the position that rational thought occupies in thinking the structure of the world, then it is plausible that this thought faces, as such, a general positional dilemma. The dilemma is that, in order to verify that it is so much as coherent, it must have recourse to knowledge that it cannot represent itself to itself as capable of establishing by its own means. The rational thought of structure must, then, either have recourse to what are essentially extra-rational (for instance intuitive) means of knowledge about the infinite, or content itself with the consequence of the inherent undecidability of its own attempt to extend its structure to the comprehension of the world as such.

If this dilemma is indeed general, it does not turn essentially on any of the specific features of Carnap's logical-syntax project on which commentators have focused, and which subsequent analytic-philosophical positions have learned to overcome. For example, it does not arise specifically from his conventionalism, or from the specific demand that logical syntax be empty of content, or from the specific aim of his project for the construction of multiple linguistic calculi in accordance with the methodical 'Principle of Tolerance'. Indeed, it appears to characterize *any* project that attempts to demonstrate the entailment of the totality of facts of the world from some more restricted (but finitely comprehensible) set of facts or truths about its structure. In this final section, I shall consider briefly how the dilemma might be formulated, and what kind of problem it might pose, for two recent projects of this general form.

In his 2012 book, *Constructing the World*, David Chalmers proposes and details a project, explicitly analogous to Carnap's own project of structural explication in the *Aufbau*, of accounting for the totality of truths on the basis of a more restricted 'compact' class of basic truths. More specifically, the totality of truths is argued to be *scrutable* from a narrower subset of them, where a truth S is scrutable from a class C, roughly, if a subject of a particular kind who knew all of the truths in class C would be in a position to know the truth S (Chalmers 2012: 40).[9] Further, a subject is said to be 'in a position to know' a truth when it is possible that the subject could come to know that truth from the subject's current position and without gaining any further empirical information (2012: 49). Various possibilities for this subset (the 'scrutability base') are considered, but most centrally Chalmers argues that a sufficient base could plausibly consist in a collection of physical truths, qualitative or phenomenal truths, certain indexical truths, and finally a 'that's all' truth specifying that the world is a *minimal* world with respect to those (other) truths (2012: 108–12).

Somewhat similarly, in his 2011 *Writing the Book of the World*, Theodore Sider argues for the possibility of giving a 'fundamental' description of the world by describing its metaphysically underlying structure. In particular, Sider argues, metaphysical inquiry can elicit a determination of a privileged class of predicates and other expressions (including logical ones) which 'carve at the joints' of reality's 'true' or most 'fundamental' structure, so that the determination and clarification of these expressions amounts to a way of 'figuring out the right categories for describing the world' (Sider 2011: 1). The analysis of claims and expressions in ordinary language will then take the form of a reductive 'metaphysical semantics' which will show how these claims and expressions in general can be reduced to those couched in the privileged language of the 'joint-carving' ones, those which capture in a privileged way the underlying structure of the world as a whole.

In Chalmers' account, the subject who is relevant to the assessment of the reducibility thesis is idealized in various ways. For instance, she possesses 'any concept that it is possible to possess', is able to entertain arbitrarily complex thoughts (possibly, Chalmers says, including infinite conjunctions of finite thoughts), and is capable of reasoning that is idealized in permitting rational calculations and proofs with arbitrarily many steps. Further, she never makes mistakes in reasoning, is sensitive to all relevant reasons for judgment, and can use 'any possible reasoning processes, regardless of whether humans actually use those processes' (Chalmers 2012: 63). All of these idealizations are meant to be extensions to the infinite of capacities and conceptual possessions that are already exercised and possessed in a more limited way by ordinary (finite) subjects. At the same time, however, in order to underwrite the scrutability thesis and thus the (idealized) knowability of *all* truths, they will require not only that these capacities be extended from the finite to the infinite, but also that the extension be *complete*: that the envisaged procedures extend, in other words, to the knowing the totality of the world or to all that is the case.

As we have seen, however, in light of the considerations that Gödel raises against the coherence of Carnap's logical-syntax project, there is an important and general reflexive problem for any such specification, grounded in the essential positional features of the reflexive moment by which rational thinking articulates its own procedures. The problem is that any such specification of procedures will evince infinitely many statements that are undecidable by those procedures, making it impossible coherently to see (from *any* perspective) the envisaged procedures as embodying principles that are simultaneously both consistent and complete in their extension to all that is the case. At any rate, it is clear at this point how the dilemma that we have discussed threatens the general claim of reducibility, or of the rational knowability of the totality of truths given only knowledge of the base class: either what is known in knowing the base class is already inconsistent, or it is not rationally extendible (by means of *any* coherently specifiable procedures, finitary *or* infinitary) to knowledge of the totality of truths.

Chalmers discusses considerations arising from Gödel's theorems in one section of *Constructing the World* specifically addressing mathematical truths as one species of 'hard cases' for the success of the scrutability thesis. As he puts the issue in the section, the challenge for the scrutability of all truths from the basis consisting only in physical, phenomenal, and indexical truths that is posed most centrally by Gödelian considerations is the possibility that, in light of Gödel's theorems, there are some (mathematical, as well as second-order logical) truths that are not rationally knowable (i.e., are not knowable by any reasoner) *a priori*; and are, thus, not knowable by any agent (no matter how idealized) who knows only the truths in the scrutability base. In light of the positional issues we have discussed, it is not clear that this is the right way to put the problem posed by the Gödelian considerations: it is plausible, in particular, that the problem is not that of whether there are truths that are not scrutable from any position, but rather that of whether there is any (one) position from which all truths (of the relevant kind) are scrutable. Leaving that aside, however, Chalmers argues that it is possible to vindicate the scrutability thesis by 'idealizing away from' our capacity to consider only finitely many cases at once, or from the requirement that proofs have only finitely many premises. In particular, Chalmers suggests, we can do so by appealing to just the kind of infinitary consequence 'rule' that Carnap himself appealed to, as we saw, in defending the claim of the overall consistency of specific syntactically formulated languages. As Carnap already suggested and Chalmers notes, if this sort of 'rule' can be appealed to, it will be possible by means of it to settle the truth-value of every statement of arithmetic (albeit not within a *single* finitely specifiable language), and a kind of completeness will apparently follow.

Given this, the question of the coherence of the envisaged perspective, and hence of the scrutability of all truths of the relevant kind, becomes that of the extent to which it is coherent to see the 'procedure' of a purportedly infinitary reasoner in applying such a rule as (something that is really coherently understandable as) a 'procedure' at all. But as we saw above, there are serious doubts to be raised here about the extent to which we can indeed see such an envisaged thinker as engaging rational procedures at all. In particular, it is not clear that we can credit any agent with the ability to carry out such a 'procedure', unless we are in a position already to credit her with knowledge of *all* properties of *all* numbers (and hence all arithmetical truths) already; but to do so would obviously be question-begging in this context. Moreover, even if there is some idealized knower who can use such infinitary rules to arrive at arbitrary truths, she will not be able to express the methodical basis of her results in any (single, consistent) language that we can understand.

But even if the relevant idealization is, after all, coherent with respect to arithmetic truths, there will still be important perspectival and positional issues of a more general kind arising from the Gödelian considerations, and plausibly bearing against *any* attempt to describe the totality of truths as consequences of a comprehensibly limited subset thereof. For as we have seen, if any such basis is to be seen as (so much as) consistent, this insight *itself* must go beyond anything that is rationally inferable from it, and so the envisaged base is then, by light of that very consideration, insufficient to establish *all* the relevant truths. If this consideration can indeed be generalized, then it appears to have the consequence that *any* specifiable scrutability base that an agent can know as such is then known to be either inconsistent (in which case it is hardly useful as a base at all) or incomplete, and the overall scrutability thesis fails.

Something similar appears to be the case with respect, as well, to Sider's project of determining the structure of the world as a matter of the 'metaphysical semantics' of 'fundamental' structure. Although Sider is much less explicit than Chalmers, in general, about the positional commitments of the project or its claims about the possibility of *knowing* truths in general on the basis of known truths about structure, it is clear that the claims that he does make for the *completeness* of the fundamental truths invite similar problems. For example, Sider holds (in a preliminary formulation) that 'every nonfundamental truth holds in virtue of some fundamental truth' and (in a more precise one) that 'every sentence that contains expressions that do not carve at the joints' can be analyzed, in principle, into sentences in terms purely of the 'metaphysical semantics' of expressions that do (2011: 105, 115). Sider says that the analysis need not yield a reduction to purely syntactic notions (2011: 113), and it need not yield explicit definitions or match intuitive judgments of cognitive significance (2011: 117). But it is clear that for the project to work, it must nevertheless be possible in principle to see the totality of truths (or sentences) as grounded (in principle) by that comprehensible (thinkable) subset of them which are expressible purely in structural or 'joint-carving' terms. And given this, it appears evident that the kind of general positional consideration that we have discussed here will apply. In particular: given any coherent conception of the 'fundamental' truths overall, we can pose the question of how and whether the way in which these truths are said to ground *all* truths is itself included in those truths. In view of the Gödelian positional considerations, it appears likely that it *cannot*, on pain of contradiction, coherently be seen as so included. It is then hard to see how the constructivist project can succeed, except by simply ignoring this kind of positional consideration: ignoring, that is, the problem of the *position* of rational thinking itself, in relation to the totality of what it can think.

If, on the other hand, the positional issue is taken seriously, and as bearing in an important way on the question of the space of possibilities open to contemporary thought, what general conclusions can we draw? As we have seen, Gödel's argument against Carnap plausibly witnesses a general and comprehensive dilemma for rational thought's understanding of itself in application to the totality of the world. In view of these considerations, the rational thought of finite beings is not apparently in a position to secure its own applicability to the world as such by means of a discernment of its 'fundamental' or underlying structure.[10] However, if the considerations explored here thus plausibly demonstrate the failure of the constructivist attempt to understand the thinking-being relationship in structuralist terms, the phenomenon of undecidability which lies at their root may nevertheless bear important implications for any contemporary understanding of the overall form of this relationship itself. For if it appears, as we have seen, to verify the essential incompleteness of any purportedly overall discernment of the world's structure, it simultaneously appears to indicate the unlimited possibility for thought ever again to revolutionize itself in its recurrently inexhaustible reflection of it.

Notes

1 For a variety of these contemporary 'structuralist' projects, see Livingston (2008: chap. 1).
2 That is, those that either have a finite number of axioms and derivation rules or are, at least, 'axiomatizable' in that the axiomhood of arbitrary sentences is (finitely) decidable. For details of the formalist project, see, e.g., Hilbert (1925).
3 This statement in fact requires the slightly strengthened version of Gödel's first theorem that is due to Rosser in 1936.
4 See, especially, Gödel's footnote 23.
5 It is worth noting that the situation, on this formulation, resembles that which plausibly leads to one of the main concerns of the later Wittgenstein's investigation of the problem of rules and their application in Wittgenstein (1951: sec. 201): namely that since no rule can apparently determine its own application, each application seems to require a new interpretation, leading to an apparent infinite regress.
6 For this kind of 'Platonist' suggestion, see, in particular, Gödel (1961).
7 It is this consequence of Gödel's views of the relationship of mathematical thought to its linguistic formulation that confirms most directly his adherence to what I called (in Livingston (2012)) the 'generic' orientation of thought (by contrast with Carnap's thoroughgoing adherence to the 'constructivist' orientation): for the four orientations of thought, see, e.g., Livingston (2012: chap. 1); for Carnap's constructivism, see (2012: chap. 2); and for Gödel as a 'generic' thinker, see (2012: 328).
8 I am indebted to John Bova for some of the considerations in this paragraph.
9 There are various formulations of slightly different scrutability relations (inferential scrutability, conditional scrutability, and a priori scrutability), but I will ignore these differences in what follows.
10 Indeed, in view of the generality of the conclusion, it seems as if something similar will apply to many other current attempts to characterize 'fundamental' notions and truths in application to the totality of facts, including such notions as those of privileged 'natural' features and properties, 'genuine features', 'intrinsic properties', 'metaphysical furniture', as well as related ideas about privileged ways of connecting linguistic terms to entities, properties, and features. In each such case, we can ask whether a coherent specification of the relevantly fundamental truths would include the basis for applying that set of truths to the totality, and in light of the Gödelian considerations in each case it seems it could not.

References

Carnap, R. (1934) *The Logical Syntax of Language*, (trans.) A. Smeaton, Paterson, NJ: Littlefield, Adams & Co., 1959.

Chalmers, D. J. (2012) *Constructing the World*, Oxford: Oxford University Press.
Gödel, K. (1951) Some Basic Theorems on the Foundations of Mathematics and their Implications, in S. Feferman (ed.), *Kurt Gödel: Collected Works*, Vol. III, New York: Oxford University Press.
Gödel, K. (1953) Is Mathematics Syntax of Language? Version III, in S. Feferman (ed.), *Kurt Gödel: Collected Works*, Vol. III, New York: Oxford University Press.
Gödel, K. (1959) Is Mathematics Syntax of Language? Version V, in S. Feferman (ed.), *Kurt Gödel: Collected Works*, Vol. III, New York: Oxford University Press.
Gödel, K. (1961) The Modern Development of the Foundations of Mathematics in the Light of Philosophy, in S. Feferman (ed.), *Kurt Gödel: Collected Works*, Vol. III, New York: Oxford University Press.
Goldfarb, W. (1995) Introductory Note to Gödel's "Is Mathematics Syntax of Language?", in S. Feferman (ed.), *Kurt Gödel: Collected Works*, Vol. III, New York: Oxford University Press.
Hilbert, D. (1925) On the Infinite, in J. van Heijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic*, Cambridge, MA: Harvard University Press.
Livingston, P. M. (2008) *Philosophy and the Vision of Language*, New York: Routledge.
Livingston, P. M. (2012) *The Politics of Logic: Badiou, Wittgenstein, and the Consequences of Formalism*, New York: Routledge.
Sider, T. (2011) *Writing the Book of The World*, Oxford: Clarendon.
Tarski, A. (1937) The Concept of Truth in Formalized Languages, in J. H. Woodger (trans.), *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*, 2nd ed., Indianapolis, IN: Hackett.
Wittgenstein, L. (1921) *Tractatus Logico-Philosophicus*, (trans.) D. F. Pears and B. F. McGuinness, New York: Routledge and Kegan Paul.
Wittgenstein, L. (1953) *Philosophical Investigations*, (trans.) G. E. M. Anscombe, P. M. S. Hacker and J. Schulte, Revised 4th ed., Oxford: Wiley-Blackwell.