Is Mathematical Insight Algorithmic ?

Martin Davis
Courant Institute of Mathematical Sciences
New York University

March 22, 1995

Roger Penrose replies "no," and bases much of his case on Gödel's incompleteness theorem: it is *insight* that enables us to *see* that the Gödel sentence, undecidable in a given formal system is actually true; how could this *insight* possibly be the result of an algorithm ? This seemingly persuasive argument is deeply flawed. To see why will require looking at Gödel's theorem at a somewhat more microscopic level than Penrose permits himself.

It will be helpful (though not essential to our argument) to place the discussion in terms of what is usually called *first order logic*. This is just the formal system which embodies the elementary classical logic of *and*, *or*, *not*, *implies*, *all*, *there exists*. In a precise formulation of first order logic, it is necessary to explain when some particular formula F is to be taken to be a *logical consequence* of a set of formulas ("premises") Γ. This can be done in two essentially different ways: *semantically* and *syntactically*. In the semantic version, F is a logical consequence of Γ if F is *true* no matter how the extra-logical symbols appearing in F and Γ are interpreted, so long as all the formulas in Γ are true under that same interpretation.

(Metaphorically: F is true in every Platonic world in which the formulas of Γ are true.) In the syntactic version, "rules of proof" involving the straightforward manipulation of symbols are specified, and F is said to be a logical consequence of Γ if F can be obtained from Γ by some finite number of applications of those rules (Penrose, p.104 gives some samples of such rules). In Gödel's 1929 doctoral dissertation, he established his famous *completeness* theorem which states that the semantic and the syntactic versions are equivalent. Moreover, this equivalence is largely independent of the detailed manner in which rules of proof are specified.

Gödel's completeness theorem answered a question Hilbert had posed in his address at the Bologna mathematical congress of 1928. Hilbert's Entscheidungsproblem for first order logic was also raised in 1928 (in the famous textbook by Hilbert and Ackermann, not at the Bologna conference as Penrose asserts), and called "the fundamental problem of mathematical logic." The problem was to give an algorithm for deciding whether a given formula was a logical consequence (in the semantic sense) of a given (finite) set of premises. Hilbert's singled out first order logic for this attention presumably because it seemed clear that all mathematical reasoning could *in principle* be carried out in this formalism.[1] For the premises one takes an *appropriate* set of mathematical axioms; a mathematical theorem is then simply a logical consequence in first order logic of those axioms. Since an

[1]It is only in this sense that Penrose is justified in somewhat loosely defining the Entscheidungsproblem as referring to "all the problems of mathematics" (p.34). Penrose's conflation of the Entscheidungsproblem with Hilbert's tenth problem of 1900, which merely asked for an algorithm for solvability of Diophantine equations, is likewise justified after the fact: it is a corollary of the methods used to give a negative solution to Hilbert's tenth problem that the question of whether any given Turing machine will eventually halt, and hence the Entscheidungsproblem, can be encoded as a Diophantine problem (Davis, Matijasevich and Robinson 1976). Of course, Hilbert in 1900 could hardly have imagined such a thing.

argument based on the rules of proof of first order logic can be checked in a completely algorithmic way, we have no trouble understanding why mathematicians should agree about proofs (p. 417) so long as they agree about the axioms (and so long as these axioms are finite in number or at least are specified by an algorithm).

In this context, Gödel's incompleteness theorem (in a strengthened form based on work of J.B. Rosser as well as the solution of Hilbert's tenth problem) may be stated as follows:

There is an algorithm which, given any *consistent* set of axioms, will output a polynomial equation $P = 0$ which in fact has no integer solutions, but such that this fact can not be deduced from the given axioms.

Here then is the true but unprovable Gödel sentence on which Penrose relies and in a particularly simple form at that. Note that the sentence is provided by an *algorithm.* If *insight* is involved, it must be in convincing oneself that the given axioms are indeed consistent, since otherwise we will have no reason to believe that the Gödel sentence is true. But here things are quite murky: great logicians (Frege, Curry, Church, Quine, Rosser) have managed to propose quite serious systems of logic which later have turned out to be inconsistent. "Insight" didn't help. New axioms are just as problematical as new physical theories, and their eventual acceptance is on not dissimilar grounds. If the underlying axioms are just the elementary axioms for the arithmetic of natural numbers ("Peano's arithmetic"), then the consistency is readily proved by unproblematical mathematical methods (which however, by Gödel's result, must go beyond what can be done with these axioms). In this case, we can indeed have confidence that the corresponding equation $P = 0$ has no solutions. If the underlying axioms are what are known as ZFC (Zermelo-Fraenkel set theory), axioms known

to be adequate for all ordinary mathematics, most mathematicians today would accept that the corresponding $P = 0$ has no solutions, but hardly on the basis of a sudden insight. Confidence in ZFC has developed slowly over almost a century. If ZFC is augmented by various of the axioms that have been proposed (e.g. axioms asserting the existence of very large infinite set or the socalled axiom of projective determinacy), most experts would be very cautious in accepting that it had been demonstrated that the corresponding $P = 0$ had no solutions.

There is certainly room for disagreement about whether the processes by which mathematical (or physical) theories are developed and accepted are algorithmic. But Gödel's theorem has nothing decisive to contribute to the discussion.

A final comment on the foundations of quantum mechanics. Penrose calls the process by which classical Hamiltonians are converted into the corresponding quantized versions by replacing real "observables" by operators as "genuine magic which works" (p.288). I have speculated elsewhere (Davis 1977) that another contribution of modern logic due to Gaise Takeuti, and ultimately going back to the seminal ideas of the Norwegian logician Thoralf Skolem, is what is needed to make sense of this "magic."

REFERENCES

Davis, Martin, Matijasevic, Yuri & Robinson, Julia (1976) Hilbert's Tenth Problem: Diophantine Equations; Positive Aspects of a Negative Solution, *Proceedings of Symposia in Pure Mathematics,* 28:323-78.

Davis, Martin (1977) A Relativity Principle in Quantum Mechanics, *International Journal of Theoretical Physics,* 16:867-74.