The Popular Impact of Gödel's Incompleteness Theorem

Torkel Franzén

Among Gödel's celebrated results in logic, there are two that can be formulated in terms that are intelligible in a general way even to those unfamiliar with the technicalities involved. The first is his completeness theorem for first order logic. This theorem, which is not widely known outside the world of logic, can be formulated as saying that every statement that follows logically from a set of axioms in a formalized language, such as that used in Zermelo-Fraenkel set theory with the axiom of choice (ZFC) or first order Peano Arithmetic (PA), can be proved using those axioms and the rules of logic. A general audience of nonmathematicians would probably find this statement of the completeness theorem intelligible but unexciting. After all, isn't that what "follows logically" means? It is no easy task to explain the distinction between the semantic concept of logical consequence and the purely formal notion of derivability so as to bring out the importance of this result for an audience unacquainted with logic or mathematics. By the time the expositor is done explaining that the proof of completeness depends essentially on the language being first order, few interested listeners or readers are likely to remain. The distinction between first order and higher order languages, although logically highly significant, is not one that holds any immediate appeal to the imagination.

The second result, Gödel's incompleteness theorem, is a very different matter. "Every sufficiently strong axiomatic theory is either incomplete or inconsistent." Many nonmathematicians at once find this fascinating and are ready to apply what they take to be the incompleteness theorem in many different contexts. The task of the expositor becomes, rather, to dampen their spirits by explaining that the theorem doesn't really apply in these contexts. But as experience shows, even the most determined wet blanket cannot prevent people from appealing to the incompleteness theorem in contexts where its relevance is at best a matter of analogy or metaphor. This is true not only of the first incompleteness theorem (as formulated above), but also of the second incompleteness theorem, about the unprovability in a consistent axiomatic theory T of a statement formalizing "T is consistent."

Supposed applications of the first incompleteness theorem in nonmathematical contexts usually disregard the fact that the theorem is a statement about formal systems and is stated in terms of mathematically defined concepts of consistency and completeness. This mathematically essential aspect is easily set aside, since "complete", "consistent", and "system" are words that are used in many different ways outside formal logic. Thus the incompleteness theorem has been invoked in justification of claims that quantum mechanics, the Bible, the philosophy of Ayn Rand, evolutionary biology, the legal system, and so on, must be incomplete or inconsistent. To dismiss such invocations of the incompleteness theorem is not to say that it doesn't make good sense to speak of these various "systems" as complete or consistent, incomplete or inconsistent. When people say that

Torkel Franzén is Universitetslektor at Luleå University of Technology in Sweden. His email address is `torkel@sm.luth.se`.

the Bible is inconsistent, they are arguing that it contains apparently irreconcilable statements, and those who regard the Bible as a complete guide to life presumably mean that they can find answers in the Bible to all questions that confront them about how to live their lives. Einstein, in regarding quantum mechanics as incomplete (although consistent), believed that it is possible to find a more fundamental physical theory. The system of laws of any country is incomplete or inconsistent or both in the sense that there are always situations in which conflicting legal arguments can be brought to bear, or in which no statute seems applicable. But the incompleteness theorem adds nothing to such claims or observations, for two reasons. The first is that these "systems" are not at all formal systems in the logical sense. There is no formally specified language, and there are no formal rules of inference in the logical sense associated with quantum mechanics, the Bible, a system of laws, and so on. What follows or does not follow from a religious or philosophical text, a scientific theory, or a system of laws is not determined by any formal rules of inference, such as might be implemented on a computer, but is very much a matter of interpretation, argument, and opinion, where the relevant reasoning is limited only by the vast boundaries of human thought in scientific, religious, political, or philosophical contexts.

The second reason for the irrelevance of Gödel's theorem in such discussions is that the incompleteness of any sufficiently strong consistent axiomatic theory established by that theorem concerns only what may be called the *arithmetical component* of the theory. A formal system has such a component if it is possible to interpret some of its statements as statements about the natural numbers, in such a way that the system proves some basic principles of arithmetic. Given this, we can produce (using Rosser's strengthening of Gödel's theorem in conjunction with the proof of the Matiyasevich-Davis-Robinson-Putnam theorem about the representability of recursively enumerable sets by Diophantine equations) a particular statement of the form "The Diophantine equation $p(x_1, \dots, x_n) = 0$ has no solution" which is undecidable in the theory, provided it is consistent. While it is mathematically a very striking fact that any sufficiently strong consistent formal system is incomplete with respect to this class of statements, it is unlikely to be thought interesting in a nonmathematical context where completeness or consistency (in some informal sense) is at issue. Nobody expects the Bible, the laws of the land, or the philosophy of Ayn Rand to settle every question in arithmetic.

There is also a different kind of appeal to the first incompleteness theorem outside mathematics, one that recognizes that the theorem applies only to formal mathematical theories. This is a line of thought that tends to appeal to postmodernists and theologians. From this point of view, the incompleteness theorem shows that even in mathematics, that supreme bastion of reason, truth is either beyond us or a matter of more or less arbitrary consensus rather than objective fact. Given our most powerful mathematical theory, we know that, assuming its consistency, we can produce an arithmetical statement such that we can add either that statement or its negation to the theory, obtaining incompatible theories that are still consistent. So either reason is powerless in this context (as in the wider context of the universe as a whole, with truth ultimately resting only in God), or there is no other truth than that which we more or less arbitrarily agree upon (just as in the physical sciences, according to this line of thought). Either way, after Gödel's theorem, mathematics flounders in a sea of undecidability.

When we look at mathematical practice, however, we find that mathematicians, although generally aware of the phenomenon of incompleteness, and therewith of the theoretical possibility that a problem they are working on may be unsolvable in the current axiomatic framework of mathematics, are by no means floundering in a sea of undecidability.

In the year 1900 David Hilbert made a famous affirmation in his presentation of twenty-three problems facing mathematics in the twentieth century [Browder 1976]. At first glance, it might be thought that the incompleteness theorem scuttles the confidence expressed in this affirmation:

> Take any definite unsolved problem, such as the question as to the irrationality of the Euler-Mascheroni constant C, or the existence of an infinite number of prime numbers of the form $2^n + 1$. However unapproachable these problems may seem to us, and however helpless we stand before them, we have, nevertheless, the firm conviction that their solution must follow by a finite number of purely logical processes. …This conviction of the solvability of every mathematical problem is a powerful incentive to the worker. We hear within us the perpetual call: There is the problem. Seek its solution. You can find it by pure reason, for in mathematics there is no *ignorabimus*.

Although Hilbert did not specify just what he meant by a "definite" problem, it is no doubt significant that his two examples are statements that can be formulated in arithmetical terms. Today, mathematicians have accepted that some apparently straightforward questions in set theory, such as the very first problem on Hilbert's list, that of the truth or falsity of Cantor's continuum hypothesis, cannot in fact be settled by a mathematical proof as proof is ordinarily understood in mathematics today. Instead, we must either rest content with proving hypothetical statements such as "Assuming CH, there is a group *G* with the properties…" or extend set theory by new axioms. Also, mathematicians who work in set theory or areas closely connected with set theory have learned to recognize the kind of problem or conjecture that may well be affected by the incompleteness of set theory. (It should be emphasized that this category of incompleteness, although established on the basis of the pioneering work in set theory by Gödel, and later Paul Cohen, is not a consequence of the incompleteness theorem.)

The situation is different with Hilbert's examples of "definite unsolved problems." It would be startling indeed if it turned out that ZFC does not settle whether or not there are infinitely many Fermat primes. In such a case, very few mathematicians would be content to note that we can consistently take the number of Fermat primes to be either finite or infinite, and leave it at that. Rather, mathematical instinct, if nothing else, tells us that whether or not there are infinitely many Fermat primes is not a question that can be meaningfully settled by stipulation, but if it can be settled at all calls for an argument that we perceive as mathematically compelling. Given such an incompleteness result, the search for new axioms in mathematics would take on a new urgency.

However, no famous arithmetical conjecture has been shown to be undecidable in ZFC. We do know that certain natural statements formalizable in arithmetic are undecidable in ZFC (given the consistency, or more accurately what is called in logic the 1-consistency, of ZFC), typically consistency statements, such as a statement formalizing "ZFC is consistent." Here again mathematical instinct tells us that whether or not ZFC is consistent cannot be meaningfully settled by stipulation, but statements of this kind are not at all what mathematicians normally seek to prove. Mathematicians tend to be content with accepting that the consistency of the most powerful formal theory to which they ordinarily refer in foundational contexts cannot be proved in ordinary mathematics, without thereby concluding that their own mathematical efforts are likely to run up against the barrier of undecidability. For, while we have no basis for a general claim that every arithmetical problem that arises naturally in mathematics is decidable in ZFC, we don't have a single example of an arithmetical problem—about primes or Diophantine equations or other such things—that has occurred to mathematicians in a natural mathematical context being shown to be unsolvable in ZFC. From a logician's point of view, it would be immensely interesting if some famous conjecture in arithmetic turned out to be undecidable in ZFC, but it seems too much to hope for. In short, while Hilbert's affirmation does not have any theoretical support from logic, logic does not refute that affirmation, as naturally understood from the point of view of the working mathematician.

It is thus understandable that the first incompleteness theorem has not had much of a "popular impact" among mathematicians, who are unlikely to seek to apply a mathematical theorem to the Bible and so on, and who are, for the reasons indicated, not overly concerned about the possibility of an arithmetical problem that they are working on being unsolvable in current mathematics. Mathematics may be "floundering" as far as solving a particular problem is concerned, but this neither leads to any inclination to regard problems such as those mentioned by Hilbert as in any way solvable by fiat or consensus, nor instills any sense that the problem may be unsolvable. This natural attitude is, furthermore, supported both by experience and by logical and philosophical argument, as briefly touched on above.

The second incompleteness theorem, although not as often referred to in nonmathematical contexts, has also prompted theologians and postmodernists to reflect that since mathematics cannot prove its own consistency, reason is powerless to justify itself, so that either there is no justification to be had, or reason can be supported only by faith. Without going into detail about such ideas, it is a relevant observation that a somewhat similar line of thought seems to have had considerable "popular impact" even among mathematicians, in their more philosophical moments. What I have in mind here is the following. Mathematicians often tend to regard proofs of consistency, not only of ZFC but of such very much weaker theories as PA, as somehow more unattainable or problematic than proofs of ordinary arithmetical statements. Indeed, it is not uncommon for mathematicians to say that arithmetic cannot be proved consistent. Thus Ian Stewart, in summing up the second incompleteness theorem for popular consumption, remarks that "Goedel showed that...if anyone finds a proof that arithmetic is consistent, then it isn't!" ([Stewart 1996], p. 262)

What is odd about such remarks is that we can easily, indeed trivially, prove PA consistent using reasoning of a kind that mathematicians otherwise use without qualms in proving theorems of arithmetic. Basically, this easy consistency proof observes that all theorems of PA are derived by valid logical reasoning from basic principles true of the natural numbers, so no contradiction is derivable in PA. It appears that many mathematicians have come to absorb the view that a consistency proof for PA is not really a consistency proof unless it convinces somebody who does not accept the axioms of PA as expressing valid principles of mathematics that PA is nevertheless consistent. The second incompleteness theorem and general experience do indeed indicate that no such proof is to be expected. If we were to make similar demands on proofs of arithmetical statements in general, we would be forced to the conclusion that it is equally impossible to prove the prime number theorem, Dirichlet's theorem, and so on. The insight underlying the idea that it is impossible to prove, in this sense, the consistency of arithmetic is a perfectly valid one, but it has nothing to do with Gödel's theorem. Instead it is the insight, familiar since antiquity, that we cannot prove everything. We need to start from some basic principles in our mathematical reasoning, principles that we can justify only in informal terms. The principles formalized in PA are the infinity of the natural number series, the basic properties of addition and multiplication, and the principle of mathematical induction. As long as we accept these principles as mathematically valid—as a large majority of mathematicians do in practice—there is no reason why we should not accept a proof of the kind described as proving the consistency of PA, just as we accept other mathematical proofs that depend on the validity of these principles. Those who do have genuine doubts about the consistency of PA will of course not accept this proof of consistency, but then there is no reason why they should accept standard proofs of the prime number theorem, Dirichlet's theorem, and so on, either.

It should be noted that in the logical literature, there are various nontrivial consistency proofs for PA, but the question of their interest and content is a subtle one, and I think it can be safely said that they will not convince anybody who has genuine doubts about the consistency of PA.

Of course, the above comments do not apply to every question of consistency. For example, the consistency of PA extended with the axiom "PA is inconsistent" is established only through the proof of the second incompleteness theorem itself, and proving the consistency of PA extended with Goldbach's conjecture as a new axiom is equivalent to proving Goldbach's conjecture. In these cases, the theory whose consistency is at issue is not one that formalizes basic principles of mathematical reasoning.

A point that deserves to be made in this connection is that the significance of consistency proofs as a means of justifying our mathematical reasoning is easily overstated. For a mathematician, it may at times seem convenient to refer to consistency in response to philosophical prodding about the truth or validity of mathematical axioms and methods of reasoning: only consistency matters, not the existence of the objects studied in mathematics or the philosophical justifiability of mathematics, and as far as we know, mathematics as it stands today is consistent. But such a view is at odds with how we actually think about arithmetical problems in mathematics. For example, there is no logical basis for claiming that there are infinitely many twin primes if all we know is that PA is consistent and proves the twin prime conjecture. Consistent theories of arithmetic may prove false theorems (when we are not talking about theorems having the logical form of Goldbach's conjecture), and if we conclude that there are infinitely many twin primes on the basis of a proof in some particular mathematical framework, the mere consistency of that framework is insufficient to justify our conclusion.

There is of course much more that could be said about the impact of the incompleteness theorem outside the field of logic proper. For one thing, there is the whole subject of Lucas-Penrose arguments in the philosophy of mind, which seek to establish that the human mind does not work on mechanical principles in mathematics by appealing to the incompleteness theorem. A more extensive discussion can be found in [Franzén 2005].

References

[1] F. Browder (ed.), *Mathematical Developments Arising from Hilbert's Problems*, American Mathematical Society, 1976. (Hilbert's address was translated into English by Mary Winston Newson for the *Bulletin of the American Mathematical Society*, 1902, and a reprint is included in this AMS volume.)

[2] T. Franzén, *Gödel's Theorem: An Incomplete Guide to its Use and Abuse*, A K Peters, 2005.

[3] Ian Stewart, *From Here to Infinity*, Oxford University Press, 1996.