

BOOLOS' NONCONSTRUCTIVE PROOF OF GÖDEL'S THEOREM

Richard E. Vesley

Gödel's incompleteness theorem has attracted many different statements and proofs. However, it seems that only recently are there appearing versions that are nonconstructive. Gödel's original presentations, e.g. in (1931), as well as succeeding generalizations and expositions, (Kleene 1950, 1952) and others, arose from or in response to the Hilbert program. They were carefully and explicitly formulated, both in statement and proof, to meet the metamathematical standards of that program, thus being neutral between classical and constructive (intuitionistic) mathematics.

Two recent nonconstructive proofs are those of Boolos and Chaitin. We discuss here the proof of Boolos (1989), amplifying our remarks originally published in (Vesley 1989). Chaitin's proof is different. Boolos and Chaitin themselves have not called attention to this aspect of their work.

As we hope to make clear below, we do not explore the nonconstructivity of such proofs in an effort to categorize them as acceptable according to some philosophy of mathematics, unacceptable according to another. Rather, we aim only to determine just what has been proved. In general, we can anticipate that, whenever a nonconstructive proof establishes a theorem, a corresponding constructive proof actually establishes a stronger theorem. This stronger theorem is often not made explicit. But here we emphasize the distinction by stating the stronger version(s).

It is expected that nonconstructive proofs of a theorem will be tidier in general than constructive proofs. Boolos' presentation of Gödel's theorem and proof fulfills this expectation, and was

introduced by Barwise on its publication (1989), as «the most straightforward proof of this result I have ever seen». The approach differs fundamentally from Gödel's in that an analog to the Berry paradox is developed, rather than an analog to the Epimenides paradox. Less easily noticed, but of concern to us here, are these differences from Gödel's results: (1) though the existence of an undecidable sentence of arithmetic is established, no effective means is given in the proof for actually producing such a sentence, and (2) the undecidable sentence asserted to exist is recognized to be true only classically and not constructively. The best way to explore these points is to state with some care precisely what is proved by Boolos' methods and what by Gödel's.

1. Boolos' proof

Suppose a *correct* algorithm M is one which never lists a false sentence of arithmetic. A *truth omitted by M* is a true sentence of arithmetic not listed by M .

Boolos' proof establishes the theorem:

(I) If M is a correct algorithm, then there is some truth omitted by M .

Proof. We say number n is *named* by formula $F(x)$ under algorithm M if M lists the formula $\forall x(F(x) \leftrightarrow x=[n])$, where $[n]$ is the numeral for number n , i.e. $[n]$ is 0 preceded by n occurrences of the symbol S (for successor).

Fixing any number k , we can choose formula $F(x)$ to express: x is the least number not named by any formula with fewer than $10k$ symbols. It is possible then to choose n so $\forall x(F(x) \leftrightarrow x=[n])$ is true. We can convince ourselves that there is such n by observing that there are only finitely many different $F(x)$ having fewer than $10k$ symbols and that each such $F(x)$ can name at most one number.

Now by choosing k large enough, we can construct $F(x)$ to have fewer than $10k$ symbols. The key is that expressing $10k$, i.e. writing $[10k]$, uses only a few more than k symbols (uses

exactly $k+13$ symbols in the standard alphabet employed by Boolos).

Finally, after choosing such large enough k and a corresponding n , let G be $\forall x(F(x) \leftrightarrow x=[n])$. Then G is a true sentence, and precisely because G is true and $F(x)$ has fewer than $10k$ symbols, G must be unlisted by M .

For more details see (Boolos 1989).

2. Gödel's proof and generalizations

Gödel in (1931) asserted:

(II) There is an algorithm M_1 such that, if M is a correct algorithm of a certain kind (provided by certain formal systems for arithmetic), then M_1 applied to M yields a (classical and constructive) truth omitted by M .

Sharper insight into algorithms started with Turing (cf. Gödel's note in (Gödel 1963: 616)) and eventually produced:

(III) There is an algorithm M_1 such that if M is a correct algorithm, then M_1 applied to M yields a (classical and constructive) truth omitted by M . (Kleene 1952: Sec. 60)

2.1. Providing an algorithm

The first important distinction is that Gödel provided, and intended to provide, an algorithm M_1 . His intention was made explicit in his statements that his proof is «constructive» and «intuitionistically unobjectionable» (1963: 609). Kleene conveys this intention when he writes «it is important that the strictly finitary metamathematical proof of Gödel's theorem should be appreciated» (1952: 206).

Boolos' proof does not establish (II) or (III) but only (I), as we show in a moment. A project remains to see whether Boolos' approach can give a constructive proof of (I), which would amount to a proof of (III) weakened to require only a classical truth omitted by M . A further project would be to adapt the approach to establish (III) itself. Is Boolos' new insight

somehow inherently nonconstructive? We note that the arithmetical least number principle, which may be needed in any version of the Berry paradox, is not in general constructively acceptable.

To define his true but unlisted sentence, Boolos uses an object $[n]$, the numeral for a certain number n , where the specification of n depends in its last step upon the observation: (1) for any given correct algorithm M and any given natural number y ($=10k$ in the argument), there is a least number n not named, according to M , by any formula with fewer than y symbols. But (1) is untenable constructively. Number n cannot be found by any algorithm (it cannot «effectively», or in Gödel's word «actually», be obtained), working from M and y . And so, since there is no algorithm for this step, Boolos' proof does not allow us to conclude that there is an algorithm by which to obtain the desired sentence.

We sketch an argument to show that no algorithm can compute the n claimed to exist by (1). Let $\exists tP(s,t)$ be any formula of arithmetic with $P(s,t)$ simply an equation between numerical terms (a formula whose truth for any particular numbers s and t can be determined by a finite numerical computation). Several famous unsolved mathematical problems (for example, the Fermat problem) have the form, for some particular P and s : does $\exists tP([s],t)$ hold? For any given natural number s , let M_s be the following algorithm. For each natural number input t , M_s computes whether the equation $P([s],[t])$ holds or not (for the fixed s); if so, M_s lists the formula: (i) $\forall x(x=0 \leftrightarrow x=0)$. (If we wish M_s to be completely defined, we can add: otherwise M_s lists $0=0$.) Then M_s lists (i) if and only if for the given s , $\exists tP([s],t)$ is true. M_s is correct, and the description of M_s enables us to describe easily a code number m_s for M_s . Note that (i) is of form $\forall x(F(x) \leftrightarrow x=0)$, for $F(x)$ containing 3 symbols.

Now suppose there is an algorithm L which, working on the code number m of any correct algorithm M and on any y , produces the least number not named, according to M , by any formula with fewer than y symbols. Apply L to m_s and 4. If the value obtained is 0, then (i) is not in the output of M_s ; if the value obtained is not 0, then 0, being less than this value, is named

according to M_s by a formula containing fewer than 4 symbols, and thus, since M_s allows no other way in which any number could be named, (i) must be in the output of M_s . In the first case, we know from the definition of M_s that $\neg\exists tP([s],t)$ holds; in the second case, that $\exists tP([s],t)$ holds. So L , if it existed, would enable us to solve in a standard way for any P and s , the problem: does $\exists tP([s],t)$ hold? To reiterate the method of solution, simply compute m_s , apply L to m_s and 4 and see whether the result is different from 0. However, not only are many unsolved problems of mathematics of this kind, it is even a theorem of recursion theory that no such solution method can exist.

2.2. Providing a constructively true omitted sentence

The second point to note about the Berry paradox argument concerns the nature of the omitted truth itself, whose existence (but not constructive existence, as we have just seen) the argument guarantees.

The Gödel diagonal argument provides an omitted true sentence of the form $\forall xA(x)$, with $A(x)$ quantifier free. Arithmetic sentences of this simple form which are true classically are also true constructively (intuitionistically). This observation is crucial in understanding the full force of the original Gödel theorem: not only does any consistent classical formal system for arithmetic fail to prove all classical truths, but also it must fail even to prove all constructive truths (a much smaller set). (And therefore of course a formal system for constructive arithmetic must also be incomplete).

Unfortunately, the Berry paradox argument, in the form given by Boolos, does not establish this. The omitted true sentence identified in this second argument, i.e. G above in section 1, is of a more complicated quantifier form, and there is no guarantee that this sentence will be constructively true.

(Some readers may be uncertain about the notion of «constructive truth» in arithmetic. If we agree that atomic truths are both classical and constructive truths, then the notion of a

more complicated truth can be defined in the classical case by the Tarski definition, in the constructive case by the Heyting definition. The latter requires the existence of algorithms at critical points, as e.g. in verifying a universal statement).

*Department of Mathematics
State University of New York at Buffalo
Buffalo, New York 14214 USA*

Bibliographical references

- BOOLOS, G. (1989). A new proof of the Gödel incompleteness theorem. *Notices of the American Mathematical Society*, 36, n° 4, 388-390.
- GÖDEL, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 37, 349-360.
- GÖDEL, K. (1963). On formally undecidable propositions of Principia Mathematica and related systems I. (english translation by van Heijenoort of (1931) with additional footnotes by Gödel). In: J. van Heijenoort (ed.), *From Frege to Gödel*. Cambridge (USA): Harvard.
- KLEENE, S.C. (1950). A symmetric form of Gödel's theorem. *Indagationes mathematicae*, 12, 244-246.
- KLEENE, S.C. (1952). *Introduction to Metamathematics*. New York & Amsterdam: van Nostrand.
- VESLEY, R. (1989). The last word on Boolos' proof. *Notices of the American Mathematical Society*, 36, n° 10, 1352.