COMPUTABILITY AND LOGIC

Third edition

GEORGE S. BOOLOS
Professor of Philosophy
Massachusetts Institute of Technology

RICHARD C. JEFFREY
Professor of Philosophy
Princeton University

CAMBRIDGE
UNIVERSITY PRESS

15
Undecidability, indefinability and incompleteness

We are now in a position to give a unified treatment of some of the central negative results of logic: Church's theorem on the undecidability of logic, Tarski's theorem on the indefinability of truth, and Gödel's first theorem on the incompleteness of systems of arithmetic. These theorems can all be seen as more or less direct consequences of the result of the last chapter, that all recursive functions are representable in Q, and a certain exceedingly ingenious lemma ('the diagonal lemma'), the idea of which is due to Gödel, and which we shall prove below. The first notion that we have to introduce is that of a ***gödel numbering***.

A ***gödel numbering*** is an assignment of natural numbers (called 'gödel numbers') to expressions (in some set) that meets these conditions: (1) different gödel numbers are assigned to different expressions: (2) it is effectively calculable what the gödel number of any expression is; (3) it is effectively decidable whether a number is the gödel number of some expression in the set, and, if so, effectively calculable which expression it is the gödel number of.

Gödel numberings enable one to regard interpreted languages supposed to be 'about' the natural numbers – i.e. having the set of natural numbers as the domain of their intended interpretation – as also referring to the numbered expressions. The possibility then arises that certain sentences, ostensibly referring to certain numbers, could be seen as referring, via the gödel numbering, to certain expressions that are *identical* with those very sentences themselves. The state of affairs just described is no mere possibility; the proof of the diagonal lemma shows how it arises, and succeeding theorems show how it may be exploited.

We shall consider a particular set of expressions and a particular gödel numbering, to which we appropriate the words 'expression' and 'gödel number'. There is nothing special about our particular gödel numbering; the theorems and proofs that we are going to give with respect to the one we use could have been given with respect to any number of others. Our expressions are finite sequences of these (distinct) symbols.

We'll make the following 'conventions' about the identity of certain symbols: we stipulate that $x_0 = x$, $x_1 = y$, $f_0^0 = \mathbf{o}$, $f_0^1 = {}'$, $f_0^2 = +$, $f_1^2 = \cdot$,

TABLE 15-1

( )	&	∃	x_0	f_0^0	f_0^1	f_0^2	…	A_0^0	A_0^1	A_0^2	…
,	v	∀	x_1	f_1^0	f_1^1	f_1^2	…	A_1^0	A_1^1	A_1^2	…
	–		x_2	f_2^0	f_2^1	f_2^2	…	A_2^0	A_2^1	A_2^2	…
	↔		.	.	.	.		.	.	.	
	→		.	.	.	.		.	.	.	
			.	.	.	.		.	.	.	

and $A_0^2 = \ =$. We now assign each symbol in Table 15-1 the number in the corresponding location in Table 15-2 as its gödel number:

TABLE 15-2

1 2	3	4	5	6	68	688	…	7	78	788	…
29	39	49	59	69	689	6889	…	79	789	7889	…
	399		599	699	6899	68899	…	799	7899	78899	…
	3999		.	.	.	.		.	.	.	
	39999		.	.	.	.		.	.	.	
			.	.	.	.		.	.	.	

We'll write 'gn' to mean 'the gödel number of'. Thus,

$$\text{gn}(x) = 5,\ \text{gn}(y) = 59,\ \text{gn}(\mathbf{o}) = 6,\ \text{gn}(') = 68,\ \text{gn}(+) = 688,$$

$$\text{gn}(\cdot) = 6889, \quad \text{and} \quad \text{gn}(=) = 788.$$

We must now extend the gödel numbering so that all finite sequences of symbols in Table 15-1 are assigned gödel numbers. (We don't distinguish between a single symbol and the sequence which consists of that one symbol.) The principle can be indicated in a single example: Since gn(∃) = 4, gn(x) = 5, gn(( ) = 1, and gn(=) = 788, we want

$$\text{gn}(\exists x(x =)$$

to be 4515788.

The principle is that if expression A has gödel number i, and B has j, then AB, the expression formed by writing A immediately before B, is to have as its gödel number the number denoted by the decimal arabic numeral formed by writing the decimal arabic numeral for i immediately before the decimal arabic numeral for j. It's clear that our gödel numbering really is a gödel numbering in the sense of the second paragraph.

of definability can be given for three- and more-place relations on natural numbers; we won't need this more general notion, however.)

A theory T is called an *extension* of theory S if S is a subset of T, i.e., if any theorem of S is a theorem of T. If f is a function that is representable in S, and T is an extension of S, then f is representable in T, and indeed is represented in T by the same formula that represents it in S. Similarly, any formula that defines a set in some theory defines it in any extension of that theory.

Lemma 3

If T is a consistent extension of Q, then the set of gödel numbers of theorems of T is not definable in T.

Proof. Let T be an extension of Q. Then diag is representable in T; for as diag is a recursive function, and all recursive functions are representable in Q, diag is representable in Q, and hence is representable in any extension of Q.

Suppose now that $C(y)$ defines the set θ of gödel numbers of theorems of T. By the diagonal lemma, there is a sentence G such that

$$\vdash_T G \leftrightarrow -C(\ulcorner G \urcorner).$$

Let $k = \text{gn}(G)$. Then

$$\vdash_T G \leftrightarrow -C(\mathbf{k}). \qquad (*)$$

Then $\vdash_T G$. For if G is not a theorem of T, then $k \notin \theta$, and so, as $C(y)$ defines θ, $\vdash_T -C(\mathbf{k})$, whence by (*), $\vdash_T G$.

So $k \in \theta$. So $\vdash_T C(\mathbf{k})$, as $C(y)$ defines θ. So, by (*), $\vdash_T -G$, and T is therefore inconsistent.

A set of expressions is called *decidable* if the set of gödel numbers of its members is a recursive set. Thus a theory T is decidable iff the set θ of gödel numbers of its theorems is recursive, iff the characteristic function of θ is recursive.

If a theory is decidable, then an effective method exists for deciding whether any given sentence is a theorem of the theory. For to determine whether a sentence is a theorem, calculate its gödel number first and then calculate the value of the (recursive, hence calculable) characteristic function for the gödel number as argument. The sentence is a theorem iff the value is 1.

Conversely, if a theory is not decidable, then *unless Church's thesis is false*, no effective method exists for deciding whether a given sentence is a theorem of the theory. For if there were such a method, then the characteristic function of the set of gödel numbers of theorems would also be effectively calculable, and hence recursive, by Church's thesis.

Theorem 1

No consistent extension of Q is decidable.

Proof. Suppose T is a consistent extension of Q. Then by Lemma 3, the set θ of gödel numbers of theorems of T is not definable in T. Now if $A(x,y)$ represented the characteristic function f of θ in T, then $A(x, \mathbf{1})$ would define θ in T. (For then if $k \in \theta, f(k) = 1$, whence $\vdash_T A(\mathbf{k}, \mathbf{1})$; and if $k \notin \theta, f(k) = 0$, whence $\vdash_T \forall y (A(\mathbf{k}, y) \leftrightarrow y = \mathbf{0})$, whence, as $\vdash_Q \mathbf{0} \neq \mathbf{1}$, $\vdash_T -A(\mathbf{k}, \mathbf{1})$.) Thus the characteristic function of θ is not representable in T, and therefore, as T is an extension of Q, not representable in Q either, and hence not recursive. So T is not decidable.

Lemma 4

Q is not decidable.

Proof. Q is a consistent extension of Q.

We can now give another proof of the proposition that first-order logic has no decision procedure, a proof that is rather different from the one given in Chapter 10.

Let L be the theory in L, the language of arithmetic, whose theorems are just the valid sentences in L. All theorems of L are theorems of Q, of course, but as not all of (indeed, none of) the axioms of Q are valid, L is not an extension of Q, and we cannot therefore apply theorem 1. But because Q has only finitely many axioms, we can nonetheless prove that L is not decidable, and hence that there is no effective method for deciding whether or not a first-order sentence is valid.

Theorem 2 (*Church's undecidability theorem*)

L is not decidable.

Proof. Let C be a conjunction of the axioms of Q. Then a sentence A is a theorem of Q iff C implies A, iff $(C \rightarrow A)$ is valid, iff $(C \rightarrow A)$ is a