GÖDEL'S THEOREM

Part of Gödel's argument was very detailed and complicated. However, it is not necessary for us to examine the intricacies of that part. The central idea, on the other hand, was simple, beautiful, and profound. This part we shall be able to appreciate. The complicated part (which also contained much ingenuity) was to show in detail how one may actually code the individual rules of procedure of the formal system, and also the use of its various axioms, into *arithmetical operations*. (It was an aspect of the profound part, though, to realize that this was a fruitful thing to do!) In order to carry out this coding we need to find some convenient way of labelling propositions with natural numbers. One way would be simply to use some kind of 'alphabetical' ordering for all the strings of symbols of the formal system for each specific length, where there is an overall ordering according to the length of the string. (Thus, the strings of length one could be alphabetically ordered, followed by the strings of length two, alphabetically ordered, followed by the strings of length three, etc.) This is called *lexicographical* ordering.* In fact Gödel originally used a more complicated numbering system, but the distinctions are not important for us. We shall be particularly concerned with *propositional functions* which are dependent on a *single* variable, like $G(w)$ above. Let the nth such propositional function (in the chosen ordering of strings of symbols), applied to w, be

$$P_n(w).$$

We can allow our numbering to be a little 'sloppy' if we wish, so that some of these expressions may not be syntactically correct. (This makes the arithmetical coding very much easier than if we

* We can think of lexicographical ordering as the ordinary ordering for natural numbers written out in 'base $k + 1$', using, for the $k + 1$ numerals, the various symbols of the formal system, together with a new 'zero', which is never used. (This last complication arises because numbers beginning with zero are the same as with this zero omitted.) A simple lexicographical ordering of strings with nine symbols is that given by the natural numbers that can be written in ordinary denary notation without zero: 1, 2, 3, 4, . . . , 8, 9, 11, 12, . . . , 19, 21, 22, . . . , 99, 111, 112, . . .

try to omit all such syntactically incorrect expressions.) If $P_n(w)$ is syntactically correct, it will be some perfectly well-defined particular arithmetical statement concerning the two natural numbers n and w. Precisely *which* arithmetical statement it is will depend on the details of the particular numbering system that has been chosen. That belongs to the complicated part of the argument and will not concern us here. The strings of propositions which constitute a *proof* of some theorem in the system can also be labelled by natural numbers using the chosen ordering scheme. Let

$$\Pi_n$$

denote the nth proof. (Again, we can use a 'sloppy numbering' whereby for some values of n the expression 'Π_n' is not syntactically correct and thus proves no theorem.)

Now consider the following propositional function, which depends on the natural number w:

$$\sim\exists x[\Pi_x \text{ proves } P_w(w)].$$

The statement in the square brackets is given partly in words, but it is a perfectly precisely defined statement. It asserts that the xth proof is actually a proof of that proposition which is $P_w(\ )$ applied to the value w itself. Outside the square bracket the negated existential quantifier serves to remove one of the variables ('there does not exist an x such that . . .'), so we end up with an arithmetical propositional function which depends on only the one variable w. The expression as a whole asserts that there is *no* proof of $P_w(w)$. I shall assume that it is framed in a syntactically correct way (even if $P_w(w)$ is not – in which case the statement would be *true*, since there can be no proof of a syntactically incorrect expression). In fact, because of the translations into arithmetic that we are supposing have been carried out, the above is actually some *arithmetical* statement concerning the natural number w (the part in square brackets being a well-defined arithmetical statement about *two* natural numbers x and w). It is not supposed to be obvious that the statement can be coded into arithmetic, but it can be. Showing that such statements can indeed be so coded is the major 'hard work' involved in the complicated part of Gödel's argument. As before, precisely *which* arithmetical statement it is will depend on the details of the numbering systems, and it will depend very much on the detailed structure of the axioms and rules of our formal system. Since all that belongs to the complicated part, the details of it will not concern us here.

We have numbered all propositional functions which depend on a single variable, so the one we have just written down must have been assigned a number. Let us write this number as k. Our propositional function is the kth one in the list. Thus

$$\sim\exists x[\Pi_x \text{ proves } P_w(w)] = P_k(w).$$

Now examine this function for the particular w-value: $w = k$. We get

$$\sim\exists x[\Pi_x \text{ proves } P_k(k)] = P_k(k).$$

The specific proposition $P_k(k)$ is a perfectly well-defined (syntactically correct) arithmetical statement. Does it have a proof within our formal system? Does its negation $\sim P_k(k)$ have a proof? The answer to both these questions must be 'no'. We can see this by examining the *meaning* underlying the Gödel procedure. Although $P_k(k)$ is just an arithmetical proposition, we have constructed it so that it asserts what has been written on the left-hand side: 'there is no proof, within the system, of the proposition $P_k(k)$'. If we have been careful in laying down our axioms and rules of procedure, and assuming that we have done our numbering right, then there cannot be any proof of this $P_k(k)$ within the system. For if there were such a proof, then the meaning of the statement that $P_k(k)$ actually asserts, namely that there is *no* proof, would be false, so $P_k(k)$ would have to be false as an arithmetical proposition. Our formal system should not be so badly constructed that it actually allows false propositions to be proved! Thus, it must be the case that there is in fact *no* proof of $P_k(k)$. But this is precisely what $P_k(k)$ is trying to tell us. What $P_k(k)$ asserts must therefore be a *true* statement, so $P_k(k)$ must be true as an arithmetical proposition. We have found a *true* proposition which has *no proof within the system*!

What about its *negation* $\sim P_k(k)$? It follows that we had also better not be able to find a proof of this either. We have just

established that $\sim P_k(k)$ must be false (since $P_k(k)$ is true), and we are not supposed to be able to prove false propositions within the system! Thus, neither $P_k(k)$ nor $\sim P_k(k)$ is provable within our formal system. This establishes Gödel's theorem.

MATHEMATICAL INSIGHT

Notice that something very remarkable has happened here. People often think of Gödel's theorem as something negative – showing the necessary limitations of formalized mathematical reasoning. No matter how comprehensive we think we have been, there will always be some propositions which escape the net. But should the particular proposition $P_k(k)$ worry us? In the course of the above argument, we have actually established that $P_k(k)$ is a *true* statement! Somehow we have managed to *see* that $P_k(k)$ is true despite the fact that it is not formally provable within the system. The strict mathematical formalists *should* indeed be worried, because by this very reasoning we have established that the formalist's notion of 'truth' must be necessarily incomplete. *Whatever* (consistent) formal system is used for arithmetic, there are statements that we can see are true but which do not get assigned the truth-value **true** by the formalist's proposed procedure, as described above. The way that a strict formalist might try to get around this would perhaps be not to talk about the concept of truth at all but merely refer to *provability* within some fixed formal system. However, this seems very limiting. One could not even frame the Gödel argument as given above, using this point of view, since essential parts of that argument make use of reasoning about what is actually true and what is not true.[2] Some formalists take a more 'pragmatic' view, claiming not to be worried by statements such as $P_k(k)$ because they are extremely complicated and uninteresting as propositions of arithmetic. Such people would assert:

> Yes, there is the odd statement, such as $P_k(k)$, for which my notion of provability or **truth** does not coincide with your instinctive notion of truth, but those statements will never come up in serious mathematics (at