Gödel's Incompleteness Theorems

Ted Sider
Philosophy of Mathematics

First incompleteness theorem Any "minimally strong" consistent axiomatic system is incomplete: for some sentence, neither it nor its negation can be proven

Second incompleteness theorem No "slightly more than minimally strong" consistent axiomatic system can prove its own consistency

1. Gödel numbering

Idea: assign to each formula a unique natural number (its "code")

Lookup table for primitive symbols in the language of arithmetic:

0	1	(	6	$\sim$	8	$\forall$	12
$'$	2	)	7	$\rightarrow$	9	$\exists$	13
$+$	3			&	10	x	14
$\times$	4			$\vee$	11	y	15
$=$	5					z	16
						$\vdots$	$\vdots$

Codes for strings: for a string of n symbols, raise the first n prime numbers to powers from the table for symbols in the string, and then multiply the results:

$$\text{code} = 2^{\overset{\text{table \# for first symbol}}{\bigcirc}} \cdot 3^{\overset{\text{table \# for second symbol}}{\bigcirc}} \cdot 5^{\overset{\text{table \# for third symbol}}{\bigcirc}} \cdot \ldots \cdot p_n^{\overset{\text{table \# for last symbol}}{\bigcirc}}$$

Example: the code of '$\forall x\, x \times 0' = x$' is:

$$2^{12} \cdot 3^{14} \cdot 5^{14} \cdot 7^{4} \cdot 11^{1} \cdot 13^{2} \cdot 17^{5} \cdot 19^{14}$$

Codes for finite sequences of strings: same, but skip the first prime number (to signal a sequence), and skip prime numbers to signal breaks between strings in the sequence

Example: the code of '$\mathsf{0} = \mathsf{0}$', '$\forall x\, x \times \mathsf{0}' = x$' is:

$$\overbrace{3^1 \cdot 5^5 \cdot 7^1}^{\mathsf{0}=\mathsf{0}} \cdot \overbrace{13^{12} \cdot 17^{14} \cdot 19^{14} \cdot 23^4 \cdot 29^1 \cdot 31^2 \cdot 37^5 \cdot 41^{14}}^{\forall x\, x \times \mathsf{0}' = x}$$

2. Arithmetic counterparts of properties

To any property, p, of strings or sequences of strings, there is a corresponding property of natural numbers.

Example: Property of strings:

> *being a string composed of exactly four symbols*

Corresponding property of natural numbers:

> *being divisible by* 2, 3, 5, *and* 7, *but not by any other prime numbers*

The numerical property is possessed by a number if and only if it's the code of a string that has the string property.

Example:. Property of sequences:

> *being a sequence of two strings*, *each of which is composed of two symbols*

Corresponding numerical property:

> *being divisible by* 3, 5, 11, *and* 13, *but not by any other prime numbers*.

More interesting properties of strings and sequences, such as the property of *Being a sequence that is a proof from certain axioms*, also have corresponding numerical properties.

3. Formalizing in the language of arithmetic

The language of arithmetic can *express* numerical counterparts of string- and sequence-properties.

Example: The string property:

being a string that contains exactly one symbol.

Corresponds to the numerical property

being divisible by 2 but not by any other prime number

which is expressed by the formula:

$$\sim x = 0 \mathbin{\&} \sim x = 0' \mathbin{\&} \forall y\big((\sim y = 0' \mathbin{\&} \exists z\ x = y \times z) \to \exists z\ y = 0'' \times z\big)$$

Thus in addition to its usual meaning (namely, "the number x is neither 0 nor 1, and any number that x is divisible by is itself divisible by 2"), the formula can also be understood as talking about strings, and saying: "x is a string that contains exactly one symbol". The formula *formalizes* this property of strings.

Similarly, this formula:

$$\exists x\Big(\sim x = 0 \mathbin{\&} \sim x = 0' \mathbin{\&} \forall y\big((\sim y = 0' \mathbin{\&} \exists z\ x = y \times z) \to \exists z\ y = 0'' \times z\big)\Big)$$

formalizes the claim that *there exists at least one string containing just one symbol*. For the numerical statement it makes is true if and only if at least one string in the language of arithmetic contains just one symbol.

Suppose that, for a certain set T of axioms in the language of arithmetic, we could construct formulas "T-Proof(x)" and "Contains-contradiction(x)" that formalize the sequence-properties of *being a proof from T* and *containing some formula that is a contradiction*. Then the following formula would formalize the claim that T is consistent:

$$\sim\exists x\big(T\text{-Proof}(x) \mathbin{\&} \text{Contains-contradiction}(x)\big)$$

4. First incompleteness theorem

"Minimally strong" means:

The language of T includes, at least, the language of arithmetic.

There is an algorithm for telling what formulas count as axioms of T

A certain minimal amount of arithmetic can be proven from the axioms of T.

Let A be any formula, with code n. Let "$\ulcorner A \urcorner$" stand for the numeral for the number n, i.e.:

$$\mathsf{0}\overbrace{{}'\ldots{}'}^{n \text{ of these}}$$

Gödel came up with a formula T-Prov(x) which "says in T" that x is provable in T, meaning:

If a formula A is provable in T then the formula T-Prov$(\ulcorner A \urcorner)$ is provable in T (P)

("T-Prov$(\ulcorner A \urcorner)$" is the result of beginning with the formula T-Prov(x), and changing all of the 'x's to $\ulcorner A \urcorner$s.)

He also showed how to construct a formula, G, that "says" (in T) *of itself* that it is not provable, in that this formula is a theorem of T:

$$G \leftrightarrow {\sim}T\text{-Prov}(\ulcorner G \urcorner) \qquad (*)$$

He then proved the first incompleteness theorem, by establishing these claims:

$$\text{If } T \text{ is consistent then } G \text{ is not provable in } T \qquad (1)$$

$$\text{If } T \text{ is consistent then } {\sim}G \text{ is not provable in } T \qquad (2)$$

Proof of (1):

> Suppose that T is consistent, and suppose for reductio that G is provable in T. Then by (P), the formula T-Prov$(\ulcorner G \urcorner)$ (that is, the formula in the language of arithmetic that "says" that G is provable in T) is itself provable in T. But also, since (*) is provable in T, the formula ${\sim}T$-Prov$(\ulcorner G \urcorner)$ is also provable in T, which contradicts the fact that T is consistent. Thus G is not provable in T.

Thus no matter how many axioms we choose, if they are consistent and minimally strong, there will always be some statements whose truth isn't settled by those axioms.

5. Second incompleteness theorem

To prove the second incompleteness theorem, Gödel showed how to *formalize in T* his argument for the first incompleteness theorem:

He found a formula "T-CON", which formalizes the claim that the T axioms are consistent. He then constructed this formula:

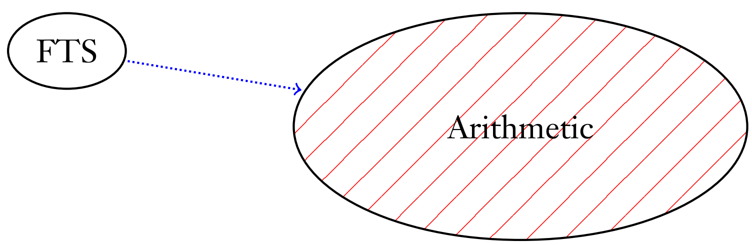
$$T\text{-CON} \rightarrow \sim T\text{-Prov}(\ulcorner G \urcorner) \tag{**}$$

This is a formalization of claim (1). He then showed how to turn the argument for (1) into a proof of (**) in T. Thus sentence (**) is provable in T.

He then argued as follows for the conclusion that if T is consistent, then T-CON isn't provable in T:

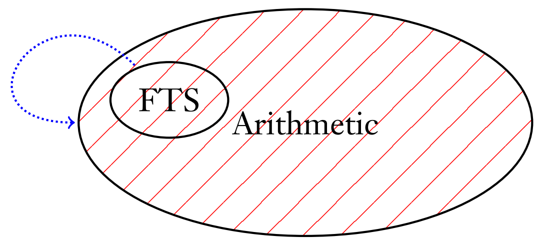
> Suppose T is consistent, and suppose for reductio that T-CON is provable in T. Then since (**) is provable in T, the formula '$\sim T$-Prov($\ulcorner G \urcorner$)' is also provable in T. And then, since (*) is provable in T, G would also be provable in T. But we showed earlier that G is *not* provable in T if T is consistent (this was claim (1)); contradiction. Therefore, T-CON is not provable in T.

The second incompleteness theorem was fatal to Hilbert's program. Hilbert was trying to use finitary methods to prove various theories, such as arithmetic, are consistent. Thus he was using a *theory*, the FTS (finitary theory of strings) to prove that arithmetic is consistent:



(The FTS ellipse represents the formulas in the finitary theory of strings; the Arithmetic ellipse represents the set of formulas in the theory of arithmetic; the diagonal hash lines indicate that there are no contradictions within the latter set; and the dotted line indicates that FTS is being used to prove that fact.)

But FTS is *part* of Arithmetic. (Hilbert is a term formalist; natural numbers are strings.). Thus the diagram should in fact look like this:



So if there are no contradictions in Arithmetic, there are none in FTS. Thus FTS could prove its own consistency. But this is impossible, given the second incompleteness theorem, since FTS satisfies the condition of only being "slightly more than minimally strong".