Gödel's Incompleteness Theorems

Romy Williamson

August 2021

Gödel's Incompleteness Theorems were first published in Gödel's 1931 paper: *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*. The paper is infamous in maths for two reasons.

1. In the context of mathematical development in the 1930's, the conclusions were counter-intuitive. Actually, the conclusions were *disturbing*. They revealed limitations in what's known as the **axiomatic method** - which had been the single thread holding all of mathematics together, ever since Euclid and his Elements. Are you alarmed yet?

2. Gödel's proof is hard. The techniques used in Gödel's paper were so original and unexpected that they seemed to come from nowhere. The 1930's mathematician-on-the street would not have been prepared for the level of subtlety involved in Gödel's argument. Even the title is bad enough.

But please read on! It isn't all bad news. Despite Gödel's original proof being fiendishly difficult to follow, you do not need to know any advanced maths to understand the main ideas. Just be prepared for a bit of mental gymnastics.

First of all, what do we mean by the **Axiomatic Method**? Basically, the job of a mathematician is to prove things. But proof by 'common sense' or 'checking the first 100 cases I can think of' are not deemed good enough - what if your intuition was wrong, or some statement failed but only after the 101st case? The only method accepted as fully watertight is to begin each proof with an agreed list of assumptions about the objects that you are working with and then show (using the rules of logic) that the hypothesis necessarily follows from the assumptions, or **axioms**.

For example, if we are working with integers then we could start by stating obvious properties of addition and multiplication e.g. $n + 0 = n$, $m + (n + 1) = (m + n) + 1$, $n \cdot 0 = 0$, $m \cdot (n + 1) = m \cdot n + m$ and then use these rules to deduce more interesting facts about integer arithmetic. In practice, it takes too long to go back to the axioms every time you write a proof, but in theory it should always be possible to fill in the gaps.

Unless the accepted rules of logic are plain wrong (in which case there's no hope for us...) then false statements can never follow from true statements. So, a perfectly careful mathematician following the axiomatic method will never accidentally prove a false statement from true axioms. Now what if the mathematician tries and tries and *can't* find a proof from the axioms for a particular statement that they think is true? First they might go for a walk, clear their head, come back and try again. It could take hours, days or even years, or it might take a different mathematician with a fresh perspective on the problem. Or, their intuition might have been mistaken and eventually someone will prove the negation (opposite) of the statement, from the axioms. But there is another possibility: what if we do not have enough axioms to allow *anyone* to prove or disprove it? We would never find out whether it is true or false from within the axiom system so we would have to either abandon our efforts, or go back and add some new axioms.

A **formal system** is a 'language' (a set of mathematical symbols, punctuation and grammar) together with a set of 'theorems' written in that language, and a set of rules for deducing new theorems from the already-proven theorems. We can set up formal systems to emulate any axiom system in maths. The difference is that when we talk about axioms, we are generally thinking of true statements about something based on reality - whereas in a formal system, we treat the theorems in the language as though they are meaningless bunches of symbols without choosing any particular interpretation.

Gödel's First Completeness Theorem talks about formal systems that are 'strong' enough to reflect a certain amount of arithmetic - i.e. to contain basic theorems about the integers (the numbers 0,1,2,3,... and their negatives). It says that for any such formal system, there is a statement in its language that can neither be proved or disproved within the formal system. In particular, we can never make a formal system that proves everything in arithmetic, by adding a non-infinite number of new axioms. Trying to shut out the undecidable statements is like trying to repair a boat with infinitely many holes - you can repair it it one place but water will always keep coming in somewhere else.

The even cooler thing about Gödel's paper is that Gödel actually constructed such an undecidable statement. To do this, Gödel invented a sort of code which is now known as Gödel Numbering. If we think of every possible statement/sequence of statements in the language as a product in a supermarket, then the Gödel number is like the serial number displayed below the bar-code. It uniquely identifies each product. The Gödel number is calculated by multiplying together some powers of prime numbers, according to which symbols appear in the statement/sequence of statements. Given any Gödel Number, it can be translated back into a statement/sequence of statements after working out its prime factorisation and making note of which prime powers appear. Pretty neat, right? Unlike computer codes such as ASCII, this code was designed with theoretical elegance in mind rather than efficiency! The Gödel number of the simplest formula, '$0 = 0$', is 243,000,000.

The purpose of this weird labelling exercise is that now we are able to make copies of statements about arithmetic, *in* arithmetic, because statements about arithmetical statements can be reframed as statements about how Gödel Numbers relate to each other. Gödel invented a symbol called **Dem** as shorthand for a relation between Gödel numbers. The statement '**Dem**(x, y)' means that "the sequence of statements that has Gödel number x is a proof for the statement that has Gödel number y". The **Dem** symbol has a complicated definition but it defines an exact, fixed, structure on the integers, therefore it belongs to arithmetic just as much as, say, the $<$ (less than) symbol. Using **Dem**, we can translate all sorts of statements into the arithmetical language, such as "the statement ... is provable".

What is the point of arithmetic 'talking about itself', you ask? Well, Gödel was ingenious enough to work out how to make a statement G with a special property. The statement G says that "the formula with Gödel number h is not provable" - the magic part being that the Gödel number of G *is* h.

So, the statement G is true if and only if G itself is not provable. Are you keeping up? Reread this last part if you need...

Now, suppose that G is false. Then G would be provable. But this is nonsense - we can't prove something that is false unless our formal system is broken/inconsistent. So, if our system is not inconsistent then G is both true, and not provable. Wow... So we have, in fact, proven an unprovable statement? This is confusing, I know. The point is that G is unprovable in our starting formal system, but we used reasoning *outside of the system* to prove it true. This reasoning was valid, but not part of the formal system itself.

So far we have quietly been assuming that the axioms we choose probably do not lead to any contradictions (consistency). Gödel's **Second** Incompleteness Theorem uses a clever argument involving our old friend G, to show whether it is possible to prove consistency from *within* a given formal system. More on this next time!