

Arithmetic is a Branch of Religion

BGGS Philosophy Café
29th July, 2004

Ken Smith

Something is rotten in the state of Denmark

Hamlet, Act I, Scene IV

I am convinced that it must be possible to find a direct proof for the compatibility of the arithmetical axioms, by means of a careful study and suitable modification of the known methods of reasoning in the theory of irrational numbers.

David Hilbert, 1900

Suppose we loosely define a *religion* as any discipline whose foundations rest on an element of faith, irrespective of any element of reason which may be present. Quantum mechanics for example would be a religion under this definition. But mathematics would hold the unique position of being the only branch of theology possessing a rigorous demonstration of the fact that it should be so classified.

Frank De Sua, 1956

Any formal system, that is interesting enough to formulate its own consistency, can prove its own consistency if, and only if, it is inconsistent.

Colloquial version of Gödel's Consistency Theorem

1 Introduction

If arithmetic is a branch of religion, then so is most of the rest of mathematics, nearly all science, economics, politics, In fact, it is difficult to think of any area of life and work in which simple arithmetic is *not* used, so that almost everything we do is religious in nature. That is, provide you accept the validity of logical reasoning — and if you don't, there are a large number of cans of very interesting worms just waiting to be opened.

David Hilbert, at the International Congress of Mathematicians held in Paris in 1900, listed 23 unsolved mathematical problems, and challenged 20th century mathematicians to provide solutions to these. Some of them have been solved, some are not yet solved. But the most interesting are those for which it has been shown that no solution exists. In the Introduction to his talk Hilbert said:

This conviction of the solvability of every mathematical problem is a powerful incentive to the worker. We hear within us the perpetual call: There is the problem. Seek its solution. You can find it by pure reason, for in mathematics there is no *ignorabimus*.

That Hilbert's faith in the power of reason was misplaced is the main topic under discussion tonight. The other topic, which will only be touched on (because the technical details are rather formidable), is the second problem in Hilbert's list (the second quotation above), the "compatibility" (or consistency, in modern usage) of the axioms of arithmetic. These were just the first two of a growing list of statements in mathematics where *ignorabimus* rules — we shall never know the solutions of some problems.

2 Completeness of Arithmetic

A logical system is described as “complete” if any proposition, which can be formulated within the system, can be proved to be either true or false, using only arguments within the system. This might seem to be a fairly trivial requirement for logical systems, but it was only in the late 19th century that it was formally recognised.

One early attempt at this was the three volume *Principia Mathematica* by Alfred North Whitehead and Bertrand Russell, published over the period 1910–1913. Very few people managed to absorb this large work, and it was not generally accepted as providing the demonstration Hilbert wanted — and it didn’t get as far as saying much about the consistency of the axioms involved.

Kurt Gödel, in a classic paper in 1931, entitled “On formally undecidable propositions in *Principia Mathematica* and related systems”, referring to Whitehead and Russell, finally put paid to any hope of simple arithmetic being a complete system.

Virtually all parts of mathematics use simple arithmetic, involving addition and multiplication of non-negative whole numbers. For simplicity, and to save space and time, in mathematics these are labelled \mathbb{N}_0 .

$$\mathbb{N}_0 = \{0, 1, 2, 3, 4, \dots\}. \quad (1)$$

Hence a proof that arithmetic is not complete implies that most of mathematics is not complete. And the result flows over to economics, physics, geology, ... and any other discipline which uses arithmetic.

2.1 Axioms of arithmetic

Since we are dealing with an infinite collection of things (the numbers listed in (1)), and we insist that all our proofs consist of a finite number of steps, we have to set out some axioms, or basic assumptions, which enable us to handle these. These axioms must also be finite in number. To give a simple illustration, the Australian Federal Budget gives figures running into billions of dollars — that is, “billion” in the US sense of 1,000,000,000 and not in the older British sense of 1,000,000,000,000. Nobody has ever counted out 1 billion dollars, \$1 at a time, then repeated the process, put the two piles together and checked that the total is \$2 billion. The Federal Treasurer relies on the rules of arithmetic in preparing the budget.

Exercise: Estimate how long it would take to count \$1 billion at a rate of \$1 per second. [Hint: there are 86,400 seconds in each day.]

There are several different ways in which axioms for arithmetic can be written down, but they all involve the concept of “successor”: this can be interpreted as “the next number”, and is usually indicated by the symbol s . In order to make things as unambiguous as possible, we make liberal use of parentheses “(” and “)”. Initially we use various symbols x, y, z, \dots but this will be varied later. First we list the axioms, with comments. Apart from the symbol \neg , which means “not”, we only use familiar bits of mathematical notation.

1. $\neg 0 = s(x)$. Zero is not the successor of any number.
2. If $s(x) = s(y)$, then $x = y$. If the successors of two numbers are the same, then the numbers themselves are the same.
3. $x + y = y + x$. This is included, since it is not true in all systems of logic.
4. $x + 0 = x$. Adding 0 doesn’t change a number.
5. $s(x + y) = x + s(y)$. Rule for finding the successor of the sum of two numbers.

6. $x \times y = y \times x$. Some students may have already encountered mathematical entities where this is not true.
7. $x \times 0 = 0$. We haven't yet defined 1 so we can't write $x \times 1 = x$, which is sometimes used.
8. $x \times s(y) = (x \times y) + x$. Rule for handling products involving the successor of a number.

These are (almost) all the axioms needed for arithmetic. It is common to include what is known as the axiom of induction, but we shall not use it. From the above we can now *define* $1 = s(0)$, and then show that $x \times 1 = x$, and *define* $2 = s(1)$ and prove that $2 \times x = x + x$, and so on. And we can define prime numbers, and show that a number can be factorised into prime factors in one and only one way (this is an essential part of Gödel's proof), and all sorts of other interesting bits of arithmetic. But first we need a bit more notation — we need enough to be able to prove theorems, and this involves the concept of the existence of numbers with certain properties, rules of logical reasoning, and a few other matters. The significance of the column headed “Code” will appear shortly.

Example: Prove that $1 + 1 = 2$, with the above definitions of 1 and 2.

$$\begin{aligned}
 1 + 1 &= s(0) + s(0), \\
 &= s(s(0) + 0) \quad \text{by Axiom 5,} \\
 &= s(s(0)) \quad \text{by Axiom 4,} \\
 &= 2.
 \end{aligned}$$

List of Symbols

| Symbol | Code | Meaning |
|--------|------|---------------------|
| 0 | 1 | zero |
| s | 2 | successor |
| + | 3 | plus |
| × | 4 | times |
| = | 5 | is equal to |
| (| 6 | left parenthesis |
|) | 7 | right parenthesis |
| , | 8 | comma |
| x | 9 | variable (a number) |
| | 10 | index for variable |
| ¬ | 11 | not |
| & | 12 | and |
| ∃ | 13 | there exists |

The usual rules of reasoning can be written down using the above symbols. We will use $F(x)$ for a formula (don't worry about the precise meaning of “formula” at this stage) involving just one thing x : this x can be a number or another formula or some other entity . . . depending on what we are discussing. This is easily extended to formulae involving more than one thing, which we can write $G(x, y)$, $H(x, y, z)$, and so on. The order of x, y, z, \dots is important. We shall not want formulae involving more than three things.

If we have two formulae F and G we need the concepts of “true” and “false” so that we can derive relations between them, and combine them in various ways. If F is true, then $\neg F$ is false — we stick to two-valued logic, or the “law of the excluded middle”. If both F and G are true we say that the combination $F \& G$ is true, while if either (or both) of them are false, then $F \& G$ is false. The existential quantifier \exists is defined by saying that $\exists x F(x)$ is true

if there is some number x for which the formula $F(x)$ is true. In the expression $\exists xG(x, y)$ we say that x is a bound variable and y is a free variable.

As examples of other logical relationships which can be derived from these, the logical “or” in F or G (which is true if either (or both) of F or G are true) can be written $\neg(\neg F \& \neg G)$. The implication $F \Rightarrow G$, read “ F implies G ”, can be written $\neg(F \& \neg G)$. If $G(x)$ is true for all x , usually written in symbols as $\forall xG(x)$, this can be written $\neg(\exists x(\neg G(x)))$.

These 13 symbols are enough for us to do everything we need to prove Gödel’s Incompleteness Theorem.

Note: Since we don’t know, before starting a proof, how many different variables we might need, we don’t use x, y, z, \dots or even x_1, x_2, x_3, \dots : we try to avoid the use of numbers as much as possible to avoid confusion — the reason for this will appear shortly. Instead we use $x|, x||, x|||, \dots$, the number of |s appearing immediately after x indicating just which variable we are referring to. This restricts our symbols to a fairly small set, but to avoid complicated notation in this exposition we will use x, y, z , and any other symbols we need. Those who wish to be pedantic can make the appropriate changes readily.

2.2 Gödel Numbers

Gödel’s main achievement (apart from the genius of discovering the theorems) was to find a way of coding all arguments and procedures into arithmetic into numbers in a unique way. Thus given a formula, or a sequence of formulae which constitute a proof, we can turn these into numbers. Hence a number can make a statement about other numbers. The essence of the proof is that Gödel found a way for a number to make a statement about itself, *without* getting involved in circular reasoning. The coding relies on the theorem of arithmetic that any number can be factorised into prime numbers in a unique manner — thus we need enough arithmetic to be able to prove this theorem.

To encourage people to follow the (slightly) complicated details, consider the formula which we will write

$$\text{Proof}(x, y, z). \tag{P}$$

Here x, y and z are all numbers. y is the Gödel number (this will be explained very shortly) of a formula with just one free variable, i.e., a variable which is not bound by the quantifier \exists . x is the Gödel number of a proof of this formula when the number z is substituted for the free variable. This forms the basis of the construction of a formula with Gödel number g which, when unscrambled, corresponds to the statement “There is no proof of the formula numbered g ”. Think about that for a moment.

Most of the remainder of the discussion will be an outline aimed at convincing you that (a) statements and proofs in arithmetic can be coded into numbers, and (b) the derivation of this formula which manages to state “There is no proof of this formula” without getting into the logical mess involving self-referential statements.

2.2.1 Coding and prime numbers

As a concrete example, take the first axiom $\neg 0 = s(x)$. From the list of symbols above the codes associated with the characters $\neg, 0, =, s, (, x$ and $)$ are, respectively, 11, 1, 5, 2, 6, 9 and 7. We cannot simply write this as 11152697 because we don’t know how to spilt up this number to regain the original formula. What Gödel did was to invoke prime numbers, and use the above sequence as the exponents of successive prime numbers, so that

$$\neg 0 = s(x) \text{ is coded as } 2^{11} \times 3^1 \times 5^5 \times 7^2 \times 11^6 \times 13^9 \times 17^7.$$

This number is known as the *Gödel number* of the first axiom. It is a rather large number, larger than 3×10^{34} , so it won’t be written out in full, but can be evaluated and factorised

using one of the available multiple precision mathematical packages. And this factorisation is unique, and represents only one formula, one of the axioms in this case. As a simpler example, $0 + 0 = 0$, with code exponents 1, 3, 1, 5, 1 is coded as the number $2^1 \times 3^3 \times 5^1 \times 7^5 \times 11^1 = 49,916,790$.

A theorem is simply a finite sequence of formulae (remember that all theorems must be finite in length), each of which is either an axiom or is derived from earlier formulae in the sequence. Suppose these formulae have Gödel numbers $G_1, G_2, G_3, \dots, G_n$ for some n . Then the Gödel number of the theorem is

$$G_T = 2^{G_1} \times 3^{G_2} \times 5^{G_3} \times \dots \times p_n^{G_n},$$

where p_n is the n^{th} prime number. If necessary, where reference is made to other theorems in the course of a proof, we may need an additional level of exponents, but nothing more complicated than this is needed.

Example: Find the Gödel number of the theorem $1 + 1 = 2$.

| Statement | Derived from | Gödel number |
|-----------------------------|------------------|--|
| $s(0) + s(0) = s(s(0) + 0)$ | from Axiom 5 | $G_1 = 2^2 \times 3^6 \times 5^1 \times 7^7 \times 11^3 \times 13^2 \times 17^6 \times 19^1 \times 23^7 \times 29^5 \times 31^2 \times 37^6 \times 41^2 \times 43^6 \times 47^1 \times 53^7 \times 59^3 \times 61^1 \times 67^7$ |
| $s(s(0) + 0) = s(s(0))$ | from Axiom 4 | $G_2 = 2^2 \times 3^6 \times 5^2 \times 7^6 \times 11^1 \times 13^7 \times 17^3 \times 19^1 \times 23^7 \times 29^5 \times 31^2 \times 37^6 \times 41^2 \times 43^6 \times 47^1 \times 53^7 \times 59^7$ |
| $s(0) + s(0) = s(s(0))$ | from above lines | $G_3 = 2^2 \times 3^6 \times 5^1 \times 7^7 \times 11^3 \times 13^2 \times 17^6 \times 19^1 \times 23^7 \times 29^5 \times 31^2 \times 37^6 \times 41^2 \times 43^6 \times 47^1 \times 53^7 \times 59^7$ |

The final statement here is the theorem we wanted to prove, and the Gödel number of the proof of the theorem is

$$G_T = 2^{G_1} \times 3^{G_2} \times 5^{G_3}, \quad (2)$$

which is a rather large but still finite, number. (For those people who have read *Mathematics and the Imagination* by Edward Kasner and James Newman, originally published in 1940, G_T here is roughly the order of magnitude of a googolplex.)

2.2.2 Checking the validity of a Gödel number

It should be clear that most numbers do not represent legitimate formulae. However the procedure for checking whether a number represents something which can be part of a formula or a sequence of formulae is straightforward, if a little involved, and can be broken down into several steps, each of which splits the number into two smaller numbers, or reduces the size of the number. Thus, after a finite (though possibly large) number of steps we are able to determine if each of the constituent parts is properly put together. Failure at any point means that we are not dealing with a formula which is part of arithmetic. The process will be illustrated by showing, in outline, how to check (a) that a number represents an integer, (b) how to check that it starts with \top , (c) how to check whether it is of the form $A \& B$, and (d) how to check whether the number represents a sequence of formulae, that is, whether it is the proof of a theorem.

1. **Question:** Is the number the Gödel number of an integer?

Answer: Is the number 2, i.e., 2^1 , which represents zero? The answer in this case is “Yes”. If not, does the number, when factorised, start with $2^2 \times 3^6$ and finish with p_n^7 ? If not, it is not an integer. If it does, remove these three factors (equivalent to removing $s($ at the start and $)$ at the end of the formula), recode the remainder by putting the exponents on smaller prime numbers and repeat the question.

2. **Question:** Is the number the Gödel number of a formula starting with \neg ?

Answer: Does the number, when factorised, start with 2^{11} ? If not, the formula (if any) is not of this type. If the answer is “Yes”, remove this factor and recode the remainder as above.

3. **Question:** Is the number the Gödel number of a formula of the form $A\&B$?

Answer: Does the number, when factorised, have an exponent 12 somewhere? If not, the formula (if any) is not of this type. If the answer is “Yes”, split the number at this point and recode each of the parts. Then ask the questions of each of the parts separately.

4. **Question:** Is the number the Gödel number of a theorem, that is, a sequence of formulae?

Answer: Does the number, when factorised, have exponents which lie outside the range 1–13, which are the code numbers associated with the symbols in the list on page 3? If not, it may be a formula, but is not a theorem. If the answer is “Yes”, then factorise the exponents, and ask the questions of each of the exponents in order.

All these procedures, and others like checking that parentheses match correctly, checking that \exists is followed by a variable which occurs later on in the formula, and so on, are purely routine and can, in principle, be carried out by a suitably programmed computer.

2.3 Gödel’s proof of the incompleteness of arithmetic

Now that we know how to code numbers, formulae and theorems, and how to check a number to determine whether or not it represents one of these, the steps in Gödel’s proof are straightforward, but need clear thinking. We start with the formula (P) given above

$$Proof(x, y, z). \tag{P}$$

z is a number, and we can check that this is so. y is the Gödel number of a formula with one free variable, and again we can check this. Then we substitute z for the free variable in y , and we can check that x represents a proof of this new formula. Thus we can check that $Proof(x, y, z)$ is a legitimate way of proceeding.

Now for the step which showed Gödel’s genius in getting around the problem of self-reference. We substitute in formula number y its own Gödel number for the free variable. This gives us the formula

$$Proof(x, y, y). \tag{Py}$$

Now consider the slightly different formula

$$\neg \exists x Proof(x, y, y). \tag{G}$$

This formula too, will have a Gödel number, which we will call g . Everything up to this point has been straightforward, and can be checked by standard procedures. (G) has one free variable, namely y .

Now consider the particular case of this formula when we substitute g , that is, its own Gödel number, for the free variable, namely

$$\neg \exists x Proof(x, g, g). \tag{Gg}$$

Before we go about trying to prove the truth or falsity of (Gg) we should understand just what such a proof would accomplish. The formula states that there is no Gödel number x which provides the proof of formula with Gödel number g , that is, a proof of formula

(G), when we substitute g for the free variable. But when we make this substitution we get formula (Gg). So, in colloquial language, formula (Gg) states “There is no proof of formula (Gg).” It would, indeed, be very surprising if such a proof existed. And it is not too hard to show that, provided arithmetic is consistent, neither formula (Gg) nor its negation can be proved. We will prove the first part of this, and make some comments about the second part.

Assume that we can prove (Gg). This means that we can find a sequence of formulae which make up the proof. This sequence will be a theorem of arithmetic, and will, therefore, have a Gödel number, m , say. This will be a proof of the formula we get from the formula with Gödel number g , that is, (G), when we substitute the number g for the free variable. If we now look back to formula (P) we see that this theorem proves the truth of (P) where $x = m$, $y = g$ and $z = g$. In other words we have proved

$$\textit{Proof}(m, g, g). \qquad \qquad \qquad (\text{Pg})$$

But our assumption was that there was no x which provided a proof of $\textit{Proof}(x, g, g)$. Since we have found one, namely m , we have a contradiction. Then if arithmetic is consistent our assumption must be wrong, and there is, indeed, no such proof.

If we can't prove (Gg) is true, can we prove it to be false? The work here is more involved, and won't be given in detail, since it relies on the rather subtle concept of ω -consistency. But the work shows that if we assume it to be false, then this, too, must be a theorem of arithmetic, and working through the gory details shows that it must be true. So again, granted the consistency of arithmetic, we can't prove that it is false.

So what Gödel managed to do was construct a valid statement of arithmetic which could not be proved to be true and could not be proved to be false. Such statements are called “undecidable”, and hence the title of his paper “On formally undecidable propositions . . .”.

But if we now look at this result from a perspective outside arithmetic, we see that Gödel produced a formula (Gg) which, when interpreted, said that there was no proof of the formula, and, in fact, showed that there was, indeed, no proof of the formula. This means that the formula makes a true statement, that is, it is true. So there is at least one true statement of arithmetic which can't be proved to be true.

3 Consistency of Arithmetic

Gödel's proof that if the axioms of arithmetic *are* consistent, then consistency cannot be proved within arithmetic is lengthy and tedious, but, in essence, does not involve any new ideas. It is far too long to give more than an outline here — and, in fact, it is rare for the full proof to be presented even in university courses dealing with mathematical logic.

It starts with the statement $0 = 1$, or, formally, $0 = s(0)$, which is obviously inconsistent with the first axiom set out above. In order to prove consistency it then formalises the argument that there is no proof of the particular statement $0 = 1$. The non-existence of such a proof is, roughly speaking, of the same nature as the earlier statement (Gg) about the non-existence of a proof. It turns out, and this is where the tedious and boring detail comes in, that if it was possible to prove the non-existence of a proof here, it would lead to a proof of the earlier non-existence proof. But since we have shown that there cannot be a proof of (Gg), then there cannot be a proof of the consistency of arithmetic, either.

So this provides an answer to Hilbert's second problem, but not the one Hilbert was looking for. It is possible to prove consistency of the axioms of arithmetic by going to a larger system, and introducing an axiom for transfinite induction. But then this larger system has the same deficiency — it is not possible to prove consistency of this larger system using only the axioms which define it.

4 Conclusion

All mathematicians are convinced that the axioms of arithmetic are consistent. But this is, in the end, a matter of belief and not of proof. So, as a number of people have remarked, mathematics is based, like religion, on faith.

And if we accept that the axioms are consistent, then we have the further paradox that, withing arithmetic, we can make a statement which cannot be proved to be true and cannot be proved to be false. But the nature of the statement is such that lack of such proofs means, looking at arithmetic from the outside, that the statement is true anyway. So we have a true statement which we cannot prove to be true.

And this isn't the end of Gödel's work in mathematical logic. The first problem Hilbert posed was about transfinite numbers, that is, a whole class of numbers starting with \aleph_0 and of increasing size. Georg Cantor posed a question about these in the 19th century and Hilbert took it up. Gödel proved that the statement was consistent with the other axioms, and Paul Cohen proved that it was independent of the other axioms. So it can be added as another axiom, or its negation can be added as an axiom, leading to two different logical systems.

And there is Gödel's work on the Axiom of Choice, and his derivation of a very peculiar solution to Einstein's field equations of general relativity.

But that is another story.

5 For Further Reading

For a Web site with a comprehensive collection of biographies of mathematicians go to <http://turnbull.mcs.st-and.ac.uk/~history/BiogIndex.html>

Weisstein, Eric W.: "Gödel's Incompleteness Theorem".

<http://mathworld.wolfram.com/GoedelsIncompletenessTheorem.html>

Barrow, John D.: *Pi in the Sky : Counting, Thinking, and Being*. Clarendon Press, 1992.

Crossley, J. N. and others: *What is mathematical logic?*. Oxford University Press, 1971.

Hofstadter, Douglas R.: *Gödel, Escher, Bach: An Eternal Golden Braid*. Penguin Books, 1980.

Smullyan, Raymond: *Forever Undecided: A Puzzle Guide to Gödel*. Oxford University Press, 1988.