Formal Systems II: Gödel's Proof

Nathan Albin

November 2, 2005

Mathematical meaning in formal systems. Remember last time we worked with the **LT**-system.

Symbols: $\{\mathbf{L}, \mathbf{T}, \mathbf{o}\}$

Axioms: $\{\mathbf{oLToo}\}$

Rules: Rule I: If $x\mathbf{LT}y$ is a theorem, then so is $x\mathbf{LTo}y$.

Rule II: If $x\mathbf{LT}y$ is a theorem, then so is $x\mathbf{oLTo}y$.

We realized that the **LT**-system contained "hidden meaning". The theorems of this system are exactly the strings that look like

$$\underbrace{\mathbf{oo...o}}_{m \text{ times}}\mathbf{LT}\underbrace{\mathbf{oo...o}}_{n \text{ times}}$$

where $1 \leq m < n$. The **LT**-system holds all information about the concept of "less than" for the positive integers. Consider the following table which summarizes the system and the interpreted meaning.

System Meaning	**Mathematical Meaning**
strings of symbols: **oooLTooooo**	mathematical statements: $3 < 5$
axiom: **oLToo**	axiom: $1 < 2$
Rule I	if $a < b$ then $a < b + 1$
Rule II	if $a < b$ then $a + 1 < b + 1$
theorems	mathematical truths

The left column represents the meaning inherent in the system itself. **LT** is a formal system as defined last time. It has symbols, strings, axioms, rules and theorems. In fact, this is all the **LT**-system really *is*. The right column represents a special way we have interpreted the system. This meaning isn't "built into" the system. Instead, it is something we observe from outside. The only reason this second meaning is at all interesting to us is because of the last row of this table. If we interpret strings of **o**s as numbers and **LT** as "less than", then it turns out that the theorems of the system correspond *exactly* to all true mathematical statements of the form $a < b$ where a and b are positive integers. If the system happened to produce the theorem **ooLTo** or **LooLT**, we would not be happy with the mathematical interpretation of the system.

Exercise: Consider the following formal system, called the **IA**-system.

Symbols: $\{\mathbf{I}, \mathbf{A}, \mathbf{o}\}$

Axioms: $\{\mathbf{ooIoAo}\}$

Rules: Rule I: If $x\mathbf{I}y\mathbf{A}z$ is a theorem, then so is $x\mathbf{I}z\mathbf{A}y$.

Rule II: If $x\mathbf{I}y\mathbf{A}z$ is a theorem, then so is $x\mathbf{oIo}y\mathbf{A}z$.

Which of the following strings are theorems of the **IA**-system? If a string is a theorem, show how you produce it. If not, what goes wrong?

(a) **oooooIoooAoo**

(b) **oooIooooAo**

(c) **oAooIooo**

Can you find a mathematical interpretation of the symbols? What do the axioms and rules "mean" mathematically?

KS-numbers. Now let's take a detour and talk about **KS**-numbers. The first thing we need to know about **KS**-numbers is that

> 31,233 is a **KS**-number.

There are also two rules which will help us find new **KS**-numbers. To understand these rules, we need to agree on some notation. Let M and N be two positive integers. By $N \operatorname{div} M$, we shall mean the result of dividing N by M and throwing away the remainder. By $N \bmod M$, we mean the remainder of the division. So

$$3 \operatorname{div} 2 = 1, \quad 3 \operatorname{div} 3 = 1, \quad 3 \operatorname{div} 1 = 3,$$
$$3 \bmod 2 = 1, \quad 3 \bmod 3 = 0, \quad 3 \bmod 1 = 0,$$
$$N \div M = N \operatorname{div} M + \frac{N \bmod M}{M}.$$

The rules for making new **KS**-numbers are

> RULE A: If N is a **KS**-number and if $(N \operatorname{div} 10^k) \bmod 100 = 12$ then
>
> $$[(N \operatorname{div} 10^k) \cdot 10 + 3] \cdot 10^k + (N \bmod 10^k)$$
>
> is also a **KS**-number.
>
> RULE B: If N is a **KS**-number and if $(N \operatorname{div} 10^k) \bmod 100 = 12$ then
>
> $$\{[(N \operatorname{div} 10^{k+2}) \cdot 10 + 3] \cdot 1,000 + 123\} \cdot 10^k + (N \bmod 10^k)$$
>
> is also a **KS**-number.

Notice that 31,233 is a **KS**-number and that

$$(31,233 \operatorname{div} 10^2) \bmod 100 = 312 \bmod 100 = 12$$

so both RULE A and RULE B apply. If we use RULE A with $N = 31,233$ and $k = 2$, we get that

$$\begin{aligned}[(31,233 \operatorname{div} 10^2) \cdot 10 + 3] \cdot 10^2 + (31,233 \bmod 10^2) \\ = [312 \cdot 10 + 3] \cdot 10^2 + 33 \\ = 3,123 \cdot 10^2 + 33 = 312,333\end{aligned}$$

is also a **KS**-number.

Similarly, if we use Rule B with $N = 31,233$ and $k = 2$, we get that

$$\begin{aligned}\{[(31,233 \operatorname{div} 10^4) \cdot 10 + 3] \cdot 1,000 + 123\} \cdot 10^2 + (31,233 \bmod 10^2) \\ = \{[3 \cdot 10 + 3] \cdot 1,000 + 123\} \cdot 10^2 + 33 \\ = \{33 \cdot 1,000 + 123\} \cdot 10^2 + 33 \\ = 33,123 \cdot 10^2 + 33 = 3,312,333\end{aligned}$$

is also a **KS**-number.

Exercise: Use the rules to make a few more **KS**-numbers. Do you see a pattern that reminds you of something we've talked about before? (Hint: Try writing the numbers without commas.)

LT-numbers. The name **KS**-numbers was perhaps a little misleading. (Of course, I didn't want to give away the surprise.) From now on, let's call them **LT**-numbers. The way we get **LT**-numbers is to start with the **LT**-system and use the trick of "Gödel numbering". To do so, we choose a number to represent each of the symbols of the **LT**-system:

$$\mathbf{L} \iff 1, \quad \mathbf{T} \iff 2, \quad \mathbf{o} \iff 3.$$

Then we can encode any string in the **LT**-system by its *Gödel number* obtained by replacing the symbols in the string by their number equivalents. For example

$$\begin{aligned} \mathbf{ooLTooo} &\iff 3,312,333 \\ \mathbf{LTLLooL} &\iff 1,211,331 \\ \mathbf{oLoLoo} &\iff 313,133. \end{aligned}$$

Note: You may assign numbers to the symbols in any way you like (You're not even restricted to single digits!) as long as you do so in such a way that no two strings give the same number. To see what can go wrong, imagine if we had said

$$\mathbf{L} \iff 1, \quad \mathbf{T} \iff 11, \quad \mathbf{o} \iff 111.$$

What string does $111,111,111$ stand for?

Notice the relationships between Rule I and Rule A and between Rule II and Rule B. In particular, see how we can use arithmetic operations to detect the presence of 12 in a number – the equivalent of detecting **LT** in a string. Multiplication and division by powers of 10 allow us to shift digits left and right. Using this with the mod operation allows us to extract portions of the number just like we can extract portions of a string. Also notice how the rules allow us to insert 3s, just like we were able to insert **o**s in the **LT**-system.

In fact, the set of **LT**-numbers exactly mirrors the set of theorems in the **LT**-system. By this, I mean that a string of **L**s, **T**s and **o**s is a theorem in the **LT**-system, if and only if its Gödel number is an **LT**-number. Here are some examples.

$$\begin{aligned} \mathbf{ooLTooo} \text{ is a theorem} &\iff 3,312,333 \text{ is an } \mathbf{LT}\text{-number} \\ \mathbf{LTLLooL} \text{ is not a theorem} &\iff 1,211,331 \text{ is not an } \mathbf{LT}\text{-number} \\ \mathbf{oLoLoo} \text{ is not a theorem} &\iff 313,133 \text{ is not an } \mathbf{LT}\text{-number}. \end{aligned}$$

A formal system for mathematics? By now we've seen how we can design formal systems which represent the concepts of "less than" and addition on positive integers. In fact, we can combine the two systems into one big system that represents both the less than comparison and the addition of positive integers.

Symbols:	$\{\mathbf{L}, \mathbf{T}, \mathbf{I}, \mathbf{A}, \mathbf{o}\}$
Axioms:	$\{\mathbf{oLToo}, \mathbf{ooIoAo}\}$
Rules:	RULE I: If $x\mathbf{LT}y$ is a theorem, then so is $x\mathbf{LTo}y$.
	RULE II: If $x\mathbf{LT}y$ is a theorem, then so is $x\mathbf{oLTo}y$.
	RULE III: If $x\mathbf{I}y\mathbf{A}z$ is a theorem, then so is $x\mathbf{I}z\mathbf{A}y$.
	RULE IV: If $x\mathbf{I}y\mathbf{A}z$ is a theorem, then so is $x\mathbf{oIo}y\mathbf{A}z$.

It seems natural to wonder if we could continue adding more symbols, axioms and rules to our system and eventually design a formal system – let's call it the $\mathbf{\Omega}$-system – powerful enough to represent all pure mathematics. In other words, we would like the $\mathbf{\Omega}$-system to...

(a) have a mathematical meaning in which the strings of symbols can be interpreted as mathematical statements.

(b) produce theorems which are always true when interpreted as mathematical statements.

(c) produce the corresponding theorem for every true mathematical statement.

What Gödel proved was that if our formal system is powerful enough to handle number theory, then we can't have all three of (a), (b) and (c). Since (a) and (b) are absolutely necessary for the formal system to be useful to us, we can interpret Gödel's result as follows. There will always be some true mathematical statements which are not theorems of the formal system.

Gödel's proof. Suppose our $\mathbf{\Omega}$-system, is powerful enough to deal with number theory. (*Principia Mathematica* is such a system.) We also want the system to have at least properties (a) and (b) from above. Can we also have (c)? As a formal system, $\mathbf{\Omega}$ will have symbols, axioms, production rules and theorems. For it to handle number theory, it's symbols must be powerful enough to enable us to encode statements like "17 is a prime number" or "9 divides 17" or "183, 231, 112 is the sum of two cubes".

One main ingredient of Gödel's proof is the use of Gödel numbering. A second ingredient requires us to have a "free variable" in our formal system. Take for example, the **LT**-system and suppose we wanted to be able to write the mathematical statement "$\mathbf{a} < 3$." To say something like this in the **LT**-system, we only need one more symbol: **a**. Let's give **a** a Gödel number, $\mathbf{a} \iff 4$. Now we can write $\mathbf{a} < 3$ in the **LT**-system if we like.

aLTooo

The Gödel number of this string is 412,333. What if we substitute this number in for **a**? Well, we get the ridiculous statement

$$\underbrace{\mathbf{oooooo..oo}}_{412{,}333 \text{ times}}\mathbf{LTooo}$$

Though it might be ridiculous in this case, Gödel used this technique of substituting a statement's own Gödel number back in for one of its variables. Using these ingredients *and* the fact that the formal system is assumed to be able to handle statements of number theory, Gödel proceeded to "break" the system. He produced a special string of symbols, lets call it G, in the formal system ($\mathbf{\Omega}$ in our case), that could be interpreted as

G is not a theorem of the $\mathbf{\Omega}$-system.

In other words, G is a string of symbols in the $\mathbf{\Omega}$-system whose mathematical interpretation is that G is not a theorem.

G is not a theorem. We can see immediately that G had better not be a theorem of the $\mathbf{\Omega}$-system. Let's pretend it is a theorem and see what happens. Since, G is a theorem, we would like to believe it because of (b). We don't want a formal system that lies to us. On the other hand, G tells us that it is NOT a theorem. This contradiction leads us to the conclusion that G is not a theorem.

G is true. On the other hand, think about the mathematical interpretation of the string of symbols called G. We saw above that G is NOT a theorem of the $\mathbf{\Omega}$-system. But this is exactly what the symbols of G mean mathematically. Therefore, there is a true mathematical statement – namely "G is not a theorem of the $\mathbf{\Omega}$-system" – which is not a theorem (when written in symbols) of the $\mathbf{\Omega}$-system.

And Gödel wins. There is no formal system for all mathematics. Any sufficiently powerful formal system which doesn't lie to us can't tell us the whole truth either.