Gödel Numbering

{x: x is a horse} is a collection that has all the worlds horses as elements, and nothing else. Thus we have

For any y, y ∈ {x: x is a horse} if and only if y is a horse.

Traveler, for example, is a horse, so Traveler ∈ {x: x is a horse}. We would expect this pattern to hold generally, so that we have:

For any y, y ∈ {x: x is ___________} if and only if y is __________,

for any way of filling in the blank.

That's what we'd expect. But now try filling in the blank with "not an element of itself"; we get:

For any y, y ∈ {x: x is not an element of itself} if and only if y is not an element of itself.

Substitute "{x: x is not an element of itself}" for "y," and we get a contradiction:

{x: x is not an element of itself} ∈ {x: x is not an element of itself} if and only if {x: x is not an element of itself} is not an element of itself.

This is *Russell's paradox*, one of several set-theoretic paradoxes that shook the foundations of mathematics around the beginning of the twentieth century. The reason these paradoxes were so disturbing is that, during the last few decades of the nineteenth century, the idea of a set had played an increasing role in clarifying and securing the foundations of mathematics, particularly the calculus. By making set theory insecure, the paradoxes threatened to undo all these gains.

A strategy for responding to the paradoxes is to adopt, hopefully on a principled basis, axioms that restrict the way the blanks can be filled in enough to prevent contradictions but not so much as to prevent set theory from playing its useful mathematical roles. That's a good questions, but how do we know that our axioms still don't still produce contradictions? The axioms are chosen to block the path by which Russell obtained a contradiction, but how can we be sure we won't still arrive at a contradiction by some, more devious path?

David Hilbert proposed an innovative solution to this problem. Ordinarily, mathematicians study things like points, lines, planes, functions, and numbers, and they use mathematical proofs as a means of learning about these things. Hilbert proposed to treat mathematical proofs as themselves the objects of mathematical investigation. A new branch of mathematics, *metamathematics*, would study proofs the way geometers study points, lines, and planes. The hope would be that an investigation of proofs would enable us to prove that the axioms of set theory wouldn't lead to a contradiction. Graph theorists investigate when it is possible to find a path from one location to another within a complex network of points connected by curves. Proof theorists do something similar, looking to see whether there is a path leading from the axioms to a contradiction. Russell's paradox has taught us to be wary of proofs in set theory. Hilbert thinks we needn't be similarly chary of the proofs produced by proof theorists. The difference is that sets are frequently infinite, and so impossible to display concretely or survey fully. Proofs, on the other hand, are finite objects that we can actually write out on paper.

Metamathematics, as Hilbert envisaged, is separate from the rest of mathematics, because it studies a different kind of thing from the things ordinary mathematicians study. Gödel devised a method of assimilating proof theory into ordinary mathematics by assigning arithmetical codes to the different symbols, thereby turning the question whether a particular sentence is provable from a certain set of axioms into an arithmetical question.

The details of the coding are fairly arbitrary. What we'll see here is one possibility among many. We'll begin by encoding terms by ordered pairs and ordered triples. For any x, y, and z, Triple(x,y,z) = Pair(x,Pair(y,z)). If w = Triple(x,y,z),then 1stin3(w) = x, 2ndin3(w) = y, and 3rdin3(w) = z.

The code for "0" is the pair Pair(1,0), which we abbreviate $\ulcorner 0 \urcorner$.

The code for "x_n" is the pair Pair(2,n), which we abbreviate $\ulcorner x_n \urcorner$.

The code for sτ is the pair Pair(4,$\ulcorner \tau \urcorner$), which we abbreviate $\ulcorner s\tau \urcorner$.

The code for $(\tau+\rho)$ is the triple Triple(5, $\ulcorner\tau\urcorner$, $\ulcorner\rho\urcorner$), which we abbreviate $\ulcorner(\tau + \rho)\urcorner$.

The code for $(\tau\bullet\rho)$ is the triple Triple(6, $\ulcorner\tau\urcorner$, $\ulcorner\rho\urcorner$), which we abbreviate $\ulcorner(\tau \bullet \rho)\urcorner$.

The code for $(\tau \text{ E } \rho)$ is the triple Triple(7, $\ulcorner\tau\urcorner$, $\ulcorner\rho\urcorner$), which we abbreviate $\ulcorner(\tau \text{ E } \rho)\urcorner$.

A number is the code of a term if and only if it is an element of every set *S* of numbers that meets the following conditions:

Pair(1,1) $\in$S.

Pair(2,i) $\in$ S.

If x is in S, so is Pair(4,x).

If x and y are in S, so are Triple(5,x,y), Triple(6,x,y), and Triple(7,x,y).

Any set that meets these conditions will be infinite, and we can't talk about infinite sets within the language of arithmetic. If we want to talk about the set of codes within the language of arithmetic, we need to figure out ways to say the things we want to say using only finite sets, so that we can take advantage of the fact that finite sets have numerical codes. Such an effort yields the following theorem:

Theorem. The set of codes of terms is a Δ set.

Proof. The set of codes of terms in Σ. x is the code of a term if and only if it's an element of a finite set s with the following properties:

If Pair(4,y) $\in$ s, y $\in$ s.

If Triple(5,y,z) $\in$ s, then y $\in$ s and z $\in$ s.

If Triple(6,y,z) $\in$ s, then y $\in$ s and z $\in$ s.

If Triple(7,y,z) $\in$ s, then y $\in$ s and z $\in$ s.

If y $\in$ s, then either y = Pair(1,0) or $(\exists n < y)$y = Pair(2,n) or

$(\exists n < y)$y = Pair(4,n) or $(\exists m < s)(\exists\, n < s)$y = Triple(5,m,n) or

$(\exists m < s)(\exists\, n < s)$y = Triple(6,m,n) or $(\exists m < s)(\exists\, n < s)$y = Triple(7,m,n).

The properties the code number of s has to satisfy for these conditions to be met can be expressed by a bounded formula.

The complement of the set of codes of terms is Σ. It will be helpful to have the following definition on board:

Definition. $x \dot{-} y = x - y$ if $x \geq y$;

$= 0$ if $x < y$.

Note that $x \dot{-} y = z$ iff $((z + y) = x \vee (x < y \wedge z = 0))$; since this is a bounded formula, $\dot{-}$ is a Δ total function.

If x isn't the code of a closed term, then, if we try to form the structure tree for x, there will a branch that doesn't terminate either in "0" or a variable. Thus, x is not a code of a term if and only if there is a finite sequence s with the following properties:

$(s)_0 = x$.

If $n < \text{length}(s)$ and $(s)_n = \text{Pair}(4,y)$, then $n+1 < \text{length}(s)$ and $(s)_{n+1} = y$.

If $n < \text{length}(s)$ and $(s)_n = \text{Triple}(5,y,z)$, then $n+1 < \text{length}(s)$ and $(s)_{n+1}$ is equal to either y or z.

If $n < \text{length}(s)$ and $(s)_n = \text{Triple}(6,y,z)$, then $n+1 < \text{length}(s)$ and $(s)_{n+1}$ is equal to either y or z.

If $n < \text{length}(s)$ and $(s)_n = \text{Triple}(7,y,z)$, then $n+1 < \text{length}(s)$ and $(s)_{n+1}$ is equal to either y or z.

If $n+1 < \text{length}(s)$, then either $(s)_n = \text{Pair}(4,(s)_{n+1})$ or $(\exists z < s)(s)_n$ is equal to either $\text{Triple}(5,(s)_{n+1},z)$ or $\text{Triple}(5,z,(s)_{n+1})$ or $\text{Triple}(6,(s)_{n+1},z)$ or $\text{Triple}(6,z,(s)_{n+1})$ or $\text{Triple}(7,(s)_{n+1},z)$ or $\text{Triple}(7,z,(s)_{n+1})$.

$(s)_{\text{length}(s) \dot{-} 1} \neq \text{Pair}(1,0)$.

$\neg(\exists n < s)(s)_{\text{length}(s) \dot{-} 1} = \text{Pair}(2,n)$.⊠

The set of codes of closed formulas is Δ; just leave off the clause for variables.

Theorem. The set of pairs <x,y> for which y is a term and x a subterm of y is Δ.

Proof: Where I say "y is a term and x is a subterm," what I really mean is that y is the code of a term and x the code of a subterm. Most of the time, I'll neglect this distinction. You'll get used to it.

If y is a term, x is a subterm of y if and only if x is an element of every finite set that meets the following conditions:

> y is in the set.
>
> If Pair(4,z) is in the set, z is in the set.
>
> If Triple(5,z,w), Triple(6,z,w), or Triple(7,zkw) is in the set, z and w are in the set.

This shows that the set of pairs <x,y> with y a term and x a subterm is Π. To see that it's also Σ, note that, for y a term. x is a subtern of y if and only if there is a number s coding a finite set with with the following properties:

> y is in the set.
>
> If Pair(4,z) is in s, z is in the set.
>
> If Triple(5,z,w), Triple(6,z,w), or Triple(7,zkw) is in tbe set, z and w are in the set.
>
> If t < s is the code of a set containing y with the properties that z is in the set coded by t whenever Pair(4,z) is in the set coded by t and that z and w are in the set coded by t whenever any of Triple(5,z,w), Triple(6,z,w), or Triple(7,z,w) in in the set coded by z, then x is an element of the set coded by t.☒

A *finite tree* is a finite set of sequences with the property that any initial segment of a member of the set is a member of the set; the statement that s is the code of a fine tree can be formalized by a bounded formula. A *finite binary tree* is a finite tree consisting entirely of sequences of 0s and 1s. Where x is a code of a term, a *structure tree* for x is a pair Pair(s,f), where s is a code of a finite binary true and f is a function with domain the set of elements of the set coded by s that satisfies the following properties:

f assigns x to the trunk of the tree, $\langle\ \rangle$.

If $y \in s$ and $f(y) = \text{Pair}(4,z)$, then $y {}^\wedge \langle 0\rangle \in s$ and $f(y {}^\wedge \langle 0\rangle) = z$, and $y {}^\wedge \langle 1\rangle \notin s$.

If $y \in s$ and $f(y) = \text{Triple}(5,z,w)$, then $y {}^\wedge \langle 0\rangle$ and $y {}^\wedge \langle 1\rangle$ are both in s, and $f(y {}^\wedge \langle 0\rangle) = z$ and $f(y {}^\wedge \langle 1\rangle) = w$.

If $y \in s$ and $f(y) = \text{Triple}(6,z,w)$, then $y {}^\wedge \langle 0\rangle$ and $y {}^\wedge \langle 1\rangle$ are both in s, and $f(y {}^\wedge \langle 0\rangle) = z$ and $f(y {}^\wedge \langle 1\rangle) = w$.

If $y \in s$ and $f(y) = \text{Triple}(7,z,w)$, then $y {}^\wedge \langle 0\rangle$ and $y {}^\wedge \langle 1\rangle$ are both in s, and $f(y {}^\wedge \langle 0\rangle) = z$ and $f(y {}^\wedge \langle 1\rangle) = w$.

If $y {}^\wedge \langle 0\rangle \in s$ and $f(y {}^\wedge \langle 0\rangle) = z$, then either $f(y) = \text{Pair}(4,z)$ or $(\exists w < s)(f(y {}^\wedge \langle 1\rangle)$ is defined and equal to w and $f(y) = \text{Triple}(5,z,w))$ or $(\exists w < s)(f(y {}^\wedge \langle 1\rangle)$ is defined and equal to w and $f(y) = \text{Triple}(6,z,w))$ or $(\exists w < s)(f(y {}^\wedge \langle 1\rangle)$ is defined and equal to w and $f(y) = \text{Triple}(7,z,w))$.

If $y {}^\wedge \langle 1\rangle \in s$ and $f(y {}^\wedge \langle 1\rangle) = w$, then, for some $z < s$, $f(y {}^\wedge \langle 0\rangle$ is defined and equal to z and either $f(y) = \text{Triple}(5,z,w)$ or $f(y) = \text{Triple}(6,z,w)$ or $f(y) = \text{Triple}(7,z,w)$.

If $y \in s$ and neither $y {}^\wedge \langle 0\rangle$ nor $y {}^\wedge \langle 1\rangle$ is in s, then either $f(y) = \text{Pair}(1,0)$ or $(\exists n < s)f(y) = \text{Pair}(2,n)$.

Unique Readability Lemma. Every (code of a) term has a unique structure tree.

The most straightforward way to prove this would be to show that the set of terms for which there is a unique structure tree contains Pair(1,1) and Pair(2,i), for each i, and that it contains Pair(4,y), Triple(5,y,z), Triple(6,y,z), and Triple(7,y,z) whenever it contains y and z. Such a proof talks about infinite set of numbers, so it can't be carried out within the language of arithmetic. To get a purely aritmetical version of the proof, we have to pick a particular number y that codes a term and show that every subterm of y has a unique structure tree, formulating the argument in such a way that it only talks about subterms ofy. I won't go through the details.

Theorem. The function Den that takes a closed term to the number it denotes is Δ.

Proof: Den(x) = v iff there is a finite set s with the following properties:

Pair(x,v) $\in$ s.

If y $\in$ s, 1st(y) is a closed term.

If Pair(Pair(1,1,), u) is in s, then y = 0.

If Pair(Pair(4,y),u) is in s, then u > 0 and Pair(y,u $\dot{-}$ 1) is in s.

If Pair(Triple(5,y,z),u) is in s, there there are t and w with Pair(y,t) and Pair(z,w) in

s

and with u = t + w.

If Pair(Triple(6,y,z),u) is in s, there there are t and w with Pair(y,t) and Pair(z,w) in

s

and with u = t · w.

If Pair(Triple(7,y,z),u) is in s, there there are t and w with Pair(y,t) and Pair(z,w) in

s

and with u = t^w.

This can be formalized by a Σ formula. Since Den is a Σ partial function with a Δ domain, it's Δ.⊠

We encode formulas the same way:

If τ and ρ are terms, the code for $\tau = \rho$ is Triple(8,$\ulcorner\tau\urcorner$,$\ulcorner\rho\urcorner$), which we abbreviate $\ulcorner\tau = \rho\urcorner$.

If τ and ρ are terms, the code for $\tau < \rho$ is Triple(9,$\ulcorner\tau\urcorner$,$\ulcorner\rho\urcorner$), which we abbreviate $\ulcorner\tau < \rho\urcorner$.

If ϕ is a formula, the code for $\sim\phi$ is Pair(10,$\ulcorner\phi\urcorner$), which we abbreviate $\ulcorner\sim\phi\urcorner$.

If ϕ and ψ are formulas, the code for $(\phi \vee \psi)$ is Triple(11,$\ulcorner\phi\urcorner$,$\ulcorner\psi\urcorner$), which we abbreviate $\ulcorner(\phi \vee \psi)\urcorner$.

If ϕ and ψ are formulas, the code for $(\phi \wedge \psi)$ is Triple(12,$\ulcorner\phi\urcorner$,$\ulcorner\psi\urcorner$), which we abbreviate $\ulcorner(\phi \wedge \psi)\urcorner$.

If ϕ and ψ are formulas, the code for $(\phi \rightarrow \psi)$ is Triple(13,$\ulcorner\phi\urcorner$,$\ulcorner\psi\urcorner$), which we abbreviate $\ulcorner(\phi \rightarrow \psi)\urcorner$.

If ϕ and ψ are formulas, the code for $(\phi \leftrightarrow \psi)$ is Triple(14,$\ulcorner\phi\urcorner$,$\ulcorner\psi\urcorner$), which we abbreviate $\ulcorner(\phi \leftrightarrow \psi)\urcorner$.

If ϕ is a formula, the code for $(\forall x_n)\phi$ is Triple(15,n,$\ulcorner\phi\urcorner$), which we abbreviate $\ulcorner(\forall x_n)\phi\urcorner$.

If ϕ is a formula, the code for $(\exists x_n)\phi$ is Triple(16,n,$\ulcorner\phi\urcorner$), which we abbreviate $\ulcorner(\exists x_n)\phi\urcorner$.

Theorem. The set of codes of formulas is Δ.

The proof is so close to the proof of the analogous proof for codes of terms that there's no real point in going through it. Same for the proof that the set of pairs $<x,y>$ with x the code of a subformula of the formula coded by y is Δ. We define structure trees for formulas the way we did for terms, and once again we have unique readability.

Theorem. The function Sub that, for θ a formula, τ a term, and n a natural number, takes $<\ulcorner\theta\urcorner,n,\ulcorner\tau\urcorner>$ to $\ulcorner\theta^{x_n}/_{\tau}\urcorner$ is Δ.

Proof: First note that the function that, given terms τ and a natural number n, takes $<\ulcorner\rho\urcorner,n,\ulcorner\tau\urcorner>$ to $\ulcorner\rho^{x_n}/_{\tau}\urcorner$ is Σ, and hence, as a Σ partial function with a Δ domain, Δ. That's so because the value the function takes with input $<\ulcorner\rho\urcorner,n,\ulcorner\tau\urcorner>$ is equal to z if and only if there is a finite set s with the following properties:

Pair($\ulcorner\rho\urcorner$,z) is in s.

If y is in s, then 1st(y) and 2nd(y) are both in s.

If Pair(Pair(1,1),y) is in s, then y = Pair(1,1,).

If Pair(Pair(2,i),y) is in s, with $i \neq n$, then y = Pair(2,i).

If Pair(Pair(2,n),y) is in s, then y = $\ulcorner\tau\urcorner$.

If Pair(Pair(4,u),y) is in s, then 1st(y) = 4 and Pair(u,2nd(y)) is in s.

If Pair(Triple(5,u,v),y) is in s, then 1st(y) = 5 and Pair(u,2ndin3(y)) and Pair(v,3rdin3(y)) are both in s.

If Pair(Triple(6,u,v),y) is in s, then 1st(y) = 6 and Pair(u,2ndin3(y)) and Pair(v,3rdin3(y)) are both in s.

If Pair(Triple(7,u,v),y) is in s, then 1st(y) = 7 and Pair(u,2ndin3(y)) and Pair(v,3rdin3(y)) are both in s.

If θ is an atomic formula, τ a term, and n a number, $\ulcorner\theta^{x_n}/_{\tau}\urcorner$ is given by:

$\ulcorner\rho = \sigma^{x_n}/_{\tau}\urcorner$ = Triple(8, $\ulcorner\rho^{x_n}/_{\tau}\urcorner$, $\ulcorner\sigma^{x_n}/_{\tau}\urcorner$).

$\ulcorner\rho < \sigma^{x_n}/_{\tau}\urcorner$ = Triple(9, $\ulcorner\rho^{x_n}/_{\tau}\urcorner$, $\ulcorner\sigma^{x_n}/_{\tau}\urcorner$).

To complete the proof, we need to do the same thing with formulas that we just did with terms. I won't give the details, which are tedious and don't involve any new ideas.⊠